

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
24 February 2005 (24.02.2005)

PCT

(10) International Publication Number
WO 2005/017129 A1

(51) International Patent Classification⁷: **C12N 1/20**,
C12Q 1/04, A61K 31/664, G01N 33/569, C07K 16/12,
14/205, C12N 15/31, 9/10, 15/54, C12P 21/02, A61K
39/106, A61P 31/04, C07F 9/24, 9/22

Ottawa, Ontario K1C 6G3 (CA). **VINOGRADOV, Evgeny**
[RU/CA]; 1541 Riverside Drive, Apt#1406, Ottawa, On-
tario K1G 4E2 (CA).

(21) International Application Number:
PCT/CA2004/001489

(74) Agent: **McKAY, Margaret**; National Research Council
of Canada, Intellectual Property Services Office, EG-12,
Bldg. M-58, 1200 Montreal Road, Ottawa, Ontario K1A
0R6 (CA).

(22) International Filing Date: 17 August 2004 (17.08.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/495,733 18 August 2003 (18.08.2003) US

(71) Applicant (for all designated States except US): **NATIONAL RESEARCH COUNCIL OF CANADA**
[CA/CA]; 1200 Montreal Road, Bldg. M-58, Room
EG-12, Ottawa, Ontario K1A 0R6 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BRISSON, Jean-Robert** [CA/CA]; 14 Simcoe Street, Ottawa, Ontario K1S
1A2 (CA). **JARRELL, Harold** [CA/CA]; 1340 Georges
Vanier Drive, Cumberland, Ontario K4C 1R6 (CA). **SZY-
MANSKI, Christine, M.** [CA/CA]; 6573 Tooney Drive,

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI,
SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: PHOSPHORAMIDE AND USES THEREOF

HS:1 [P-4)-α-D-Gal-1-3-Gro-1-]_n

NCTC11168 {6-O-Me}D-glycero-α-L-glc-Hep-1-
(HS:2) ₃
[-2)-β-D-Ribf-1-5-β-D-Gal/Nac-1-4-α-D-GlcA6{NGro}-1-]_n
₃
{OP=O(NH₂)OMe}

HS:19 [-4)-β-D-GlcA6(NGro)-1-3-β-D-GlcNac-1-]_n

HS:23/HS:36 [-3)-β-D-GlcNac-1-3-α-D-Gal-1-2-6d-α-D-altro-Hep-1-]_n
-6d-3Me-α-D-altro-Hep-1-]_n
-D-glycero-α-D-altro-Hep-1-]_n
-3Me-D-glycero-α-D-altro-Hep-1-]_n

HS:41 [-2)-β-L-Araf-1-2-β-D-6d-altro-Hepf-1-2-β-L-6d-Alt-1-]_n
-α-D-Fucf-1-]_n

(57) Abstract: There is provided herein phosphoramidate OP=O(NH₂)OMe and alkyl methyl amidophosphate and uses thereof in the identification, treatment and diagnosis of *Campylobacter jejuni*.



WO 2005/017129 A1



Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

PHOSPHORAMIDE AND USES THEREOF

FIELD OF THE INVENTION

The invention relates to cell surface structures useful in the identification and
5 targeting of *C. jejuni*.

BACKGROUND

Campylobacter jejuni is the major bacterial cause of gastrointestinal disease in
developed countries and infection can lead to the development of the neuropathy
10 known as Guillain-Barré syndrome.

Carbohydrates are implicated in a variety of functions in all domains of life.
There continues to be a growing demand for methodologies that can analyze
carbohydrate structures with increasing levels of sensitivity and simplicity.

Genome sequencing of *C. jejuni* NCTC11168 demonstrated that the strain
15 contained four gene clusters necessary for carbohydrate biosynthesis. The flagellar
modification locus, adjacent to the flagellin structural genes *flaA* and *flaB*, encodes
enzymes involved in the biosynthesis of *O*-linked pseudaminic acid and its
derivatives. The LOS and adjacent protein glycosylation loci encode enzymes
involved in the formation of outer core ganglioside mimics and bacillosamine-
20 containing *N*-linked heptasaccharide, respectively. While the capsular biosynthesis
locus, containing a Kps transport system similar to that found in other encapsulated
organisms, transfers a branched tetrasaccharide repeat to the outer membrane surface.
The CPS product of this locus has been demonstrated to be the major serodeterminant
in the heat stable typing scheme first described by Penner and Hennessy. However,
25 an inconsistency in the literature developed when only a limited number of *C. jejuni*
serotypes were believed to produce capsules based on detection by immunoblotting
yet, all strains examined contained *kps* genes necessary for capsule transport. It was
then shown in *C. jejuni* 81-176 that the high molecular weight CPS was antigenically
variable but it remained to be determined whether the loss in CPS reactivity was due
30 to the lack of CPS production or changes in its structure.

Since capsular polysaccharides are the outermost structure on the bacterial cell
they play an important role in the interaction between the pathogen, host, and
environment. In *C. jejuni* 81-176 the capsule is involved in INT407 cell invasion,

virulence in ferrets, serum resistance and maintenance of bacterial cell surface hydrophilicity.

The *C. jejuni* *pgl* locus encodes enzymes necessary for the glycosylation of multiple proteins and disruption of this pathway by mutagenesis results in multiple pleiotrophic effects. The structure of the *N*-linked glycan is GalNAc- α 1,4-GalNAc- α 1,4-[Glc- β 1,3-]GalNAc- α 1,4-GalNAc- α 1,4-GalNAc- α 1,3-Bac. However, it was unknown in the literature whether the same *N*-linked glycan was present in multiple campylobacter isolates or whether slight structural variations exist as is observed for the campylobacter *O*-linked flagellin glycan.

C. jejuni LOS have received much attention due to their unique mimicry of human ganglioside structures and their potential involvement in the induction of the autoimmune polyneuropathies, Guillain-Barré (GBS) and Miller Fisher syndromes. *C. jejuni* LOS have also recently been shown to be phase variable and important in virulence. However, in the elucidation of *C. jejuni* LOS structures there are two major problems, the need for a large amount of biomass and the time consuming effort to isolate and purify LOS.

Capsular polysaccharides (CPSs) are found on the surface of a large number of bacterial species. CPSs are known to play an important role in bacterial survival and persistence in the environment and often contribute to pathogenesis. In addition, through structural variation, the potential to mimic host cell antigens, and the ability to resist innate mechanisms such as phagocytosis and complement-mediated killing, bacterial CPSs play a role in evasion of host immune responses.

Assembly of these surface polysaccharides is remarkably conserved in bacteria. Nucleotide diphosphate sugars are synthesized in the cytoplasm and sequentially added by glycosyltransferases to an undecaprenyl pyrophosphate carrier anchored in the membrane. Many Gram-negative bacteria flip the assembled polysaccharide across the membrane using an ABC transporter consisting of the transmembrane channel, KpsM, and the ATPase, KpsT. These transporters form a complex with 4-5 additional Kps proteins to ensure proper translocation of the polysaccharide to the bacterial surface. The genetic organization of the capsule gene clusters is also conserved in bacteria with *kps* transporter genes flanking polysaccharide biosynthesis genes, an organization conducive to genetic recombination and reorganization.

Identification of *kps* genes potentially involved in capsule biosynthesis during sample sequencing of the shot-gun library of NCTC11168 prompted a systematic genetic analysis of the corresponding locus and resulted in identification of CPSs in a number of strains of *C. jejuni*. These molecules were found to be the major antigens
5 in the Penner serotyping scheme. Similar experiments performed on *C. jejuni* 81-176 confirmed these findings and demonstrated a role for the capsule in serum resistance, epithelial cell invasion and diarrhoeal disease. Subsequent characterisation of the CPSs by Alcian blue staining led to the visualisation of capsule by electron microscopy. These experiments suggested that the previously described high
10 molecular weight "lipopolysaccharides" (HMW LPSs) of *C. jejuni* are in fact CPSs.

Recently, the CPS structure of NCTC11168 was determined to contain 6-O-methyl-D-*glycero*- α -L-*gluco*-heptose, β -D-glucuronic acid modified with 2-amino-2-deoxyglycerol, β -D-GalNAc and β -D-ribose. There are several notable features encoded by the *cps* locus of NCTC11168 that correlate well with the published
15 structure: homologues of the GDP-D-*glycero*-D-*manno*heptose pathway (GmhA2, HddA and HddC); homologue of the UDP-glucose dehydrogenase, Udg, involved in the formation of UDP-glucuronic acid; and a homologue of the UDP-pyranose mutase, Glf, predicted to catalyse the reversible conversion of pyranoses to furanoses and shown to cause loss of CPS when mutated in NCTC11168 {St Michael, 2002
20 #212}.

Early studies of the structural analysis of HMW LPSs (now realised as CPSs) of *C. jejuni*, showed that these molecules are highly heterogeneous. Microarray hybridisation analysis also demonstrated some differences in the CPS-related genes between the strains of various serotypes. However, hybridisation analysis does not
25 allow detailed investigation of gene content. Sequencing of the *C. jejuni* NCTC11168 genome revealed that the GC content of the *cps* locus (*cj1415-cj1442*) is lower (26.5%) in comparison to that for the entire genome (30.6%) suggesting that this locus was acquired through horizontal gene transfer. In addition, the biosynthetic region of the *cps* locus is also prone to phase variation due to the presence of six
30 genes with homopolymeric tracts. It was subsequently shown that CPS from 81-176 undergoes antigenic variation at high frequency {Bacon, 2001 #189} and that CPS from NCTC11168 can vary in structure. However, the genetic mechanisms underlying the structural heterogeneity and antigenic variation remain unknown.

SUMMARY OF THE INVENTION

There is disclosed herein phosphoramidate structures found on the surface of a majority of *Campylobacter jejuni* isolates and uses thereof.

C. jejuni produces a capsular polysaccharide (CPS) that is the major antigenic component of the classical Penner serotyping system. High resolution magic angle spinning (HR-MAS) NMR was used to examine capsular polysaccharides directly from campylobacter cells and showed profiles similar to those observed for purified polysaccharides analysed by solution NMR. This method also exhibited the potential for campylobacter serotyping, mutant verification, and preliminary sugar analysis. HR-MAS NMR examination of growth from individual colonies of *C. jejuni* NCTC11168 indicated that the capsular glycan modifications are also phase variable. These variants show different staining patterns on deoxycholate-PAGE and reactivity with immune sera. One of the identified modifications, that showed both reduced reactivity with silver staining and rabbit sera, was a novel $-\text{OP}=\text{O}(\text{NH}_2)\text{OMe}$ phosphoramidate not observed previously in nature. This modification was attached to the 3-position of the CPS GalNAc.

Biosynthetic *cps* regions were sequenced, ranging in size from 15 to 34 kb, from *C. jejuni* strains of HS:1, HS:19, HS:23, HS:36, HS:23/36 and HS:41 serotypes and compared with the sequenced strain, NCTC11168 (HS:2). Extensive structural studies, including HR-MAS NMR, demonstrated polysaccharide heterogeneity in campylobacter CPS and demonstrated the presence of additional CPS modifications and the commonality of the recently described phosphoramidate.

Development of a novel HRMAS filtering method has allowed investigation of multiple isolates of *C. jejuni* from various clinical presentations and geographical locations and revealed that the phosphoramidate is common to approximately 70% of all strains examined. This modification appears specific to *C. jejuni* and was not observed in the closely related *Campylobacter coli*. Structural analysis of the HS:1 and HS:19 strains demonstrated that the phosphoramidate can be attached to different sugars in different linkages. Multiple phosphoramidate signals are observed during HRMAS analysis suggesting that the modification is attached to varying capsule

backbones, attached to alternate structures and/or being detected as biosynthetic intermediates.

Sequential inactivation of the *cps* biosynthetic genes in *C. jejuni* NCTC11168 followed by phosphoramidate filter analysis has allowed identification of multiple genes encoding enzymes involved in the biosynthesis of phosphoramidate: *cj1416c*, *cj1417c*, *cj1418c* and *cj1421c* (and potentially the duplicated gene, *cj1422c*). All or most of these genes are missing in other *Campylobacter* species and genome sequenced strains belonging to the epsilon proteobacteria confirming their inability to synthesize phosphoramidate. Preliminary examination of the *C. jejuni* mutants in human cell culture assays has demonstrated that the phosphoramidate is required for efficient adherence but is not necessary for cell invasion. Furthermore, the expression of this modifications renders the bacteria more sensitive to human sera.

In an embodiment of the invention there is provided use of the phosphoramidate $OP=O(NH_2)OMe$ or an immunologically active derivative thereof in the identification of *Campylobacter jejuni*. In some instances the phosphoramidate is used as a target for a binder.

In some instances, the phosphoramidate further includes an alkyl group attached to the O which is attached to the P of the methyl amidophosphate group, wherein the alkyl group is a sugar producible in *Campylobacter jejuni*. The sugar may be a naturally occurring sugar, an enantiomer, or other variant, or a non-naturally-occurring sugar.

In an embodiment of the invention there is provided use of alkyl methyl amidophosphate or an immunologically active derivative thereof in the identification of *Campylobacter jejuni*, wherein the alkyl group is a sugar producible in *Campylobacter jejuni*. The phosphoramidates discussed above may be used as vaccines in mammal to compete or reduce the severity of *C. jejuni* infection.

In an embodiment of the invention there is provided a method of modulating the adhesion of *C. jejuni* cells to a surface, the method comprising modulating the concentration of binders in surrounding fluid.

In an embodiment of the invention there is provided a substantially pure pharmaceutical composition comprising one or more the phosphoramidate described herein and a physiologically acceptable carrier. Pharmaceutical compositions of interest include those containing immunogenic conjugates and/or immunostimulants

capable of enhancing immune response in a mammal (by way of non-limiting example, *E. coli* labile toxin).

In an embodiment of the invention there is provided at least one of the phosphoramides described above which is linked through the O which is attached to the P of the methyl amidophosphate group, to an amino acid or alkyl group wherein
5 the alkyl group is a sugar producible in *Campylobacter jejuni* and the amino acid is an amino acid producible in *Campylobacter jejuni*.

In an embodiment of the invention there is provided a kit comprising:

- a) the binder as described herein; and
- 10 b) instructions for carrying out the method of claim 1 or 13.

In an embodiment of the invention there is provided use of an isolated nucleic acid encoding *Campylobacter jejuni* Cj1421c or Cj1422c or a portion or variant thereof in producing a polypeptide sequence having wild-type transferase activity in producing an amino acid sequence useful in producing non-naturally occurring antigenic
15 compounds comprising the phosphoramide of claim 16 or 17.

In some cases the nucleic acid sequence encodes an amino acid sequence at least 90% identical to the wild type Cj1421c or Cj1422c sequence.

In an embodiment of the invention there is provided use of an amino acid sequence encoding *Campylobacter jejuni* Cj1421c or Cj1422c or a portion or variant thereof having wild-type transferase activity in producing an amino acid sequence
20 useful in producing non-naturally occurring antigenic compounds comprising the phosphoramide described herein. In some instances the amino acid sequence encodes a variant at least 90% identical to the wild type Cj1421c or Cj1422c sequence. The wild type sequence of Cj1421c or Cj1422c can be readily determined by reference to published sequences. It will be understood that some variability between wildtype
25 sequences of different strains may occur and such variant wild type sequences are all "wild type" so long as they exhibit transferase activity sufficient for the production of the phosphoramide described herein. By way of non-limiting example, example, previously disclosed amino acid sequences of potential interest are listed in Table XV.
30 A transferase will be considered to have "wild type" transferase activity if it is capable of producing in vitro the phosphoramide described herein at at least 60% of the level of a wild-type transferase under the same conditions.

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 Is a depiction of Proton NMR spectra of NCTC11168 and HS:2 serostrain. (a) ^1H spectrum of NCTC11168 purified CPS with the structure of the major component shown above. The anomeric, OMe and NAc resonances are labeled. HR-MAS proton NMR spectra with 10 ms CPMG filter of NCTC11168 (b) whole cells, (c) 1/100 dilution of whole cells, and (d) HS:2 serostrain. The Asp (aspartic acid) resonances are labeled in (d). The HOD resonance at 4.8 ppm was saturated and digitally filtered affecting the intensity of the anomeric resonance C in b) and c).

FIGURE 2 Is a depiction of comparison of individual colonies of NCTC11168. (a) Silver-stained deoxycholate-PAGE: lane 1-NCTC11168 wild type population; lane 2-NCTC11168 variant #1; lane 3-NCTC11168 variant #2; lane 4-NCTC11168 variant #3. (b) Western blot of same samples loaded in same order and immunodetected with HS:2 typing sera. HR-MAS NMR spectrum of (c) the wild type population, (d) variant #1 with arrow indicating presence of an ethanolamine resonance, (e) variant #2 with arrow indicating presence of the novel modification, and (f) variant #3 with arrow indicating loss of OMe resonance. The anomeric resonances are labeled A, B, C and D. Also note the movement of the anomeric peak for residue C in all variant spectra. The HOD resonance at 4.8 ppm was saturated and digitally filtered sometimes affecting the intensity of the anomeric resonance C.

FIGURE 3 Is a depiction of NMR experiments for the *C. jejuni* NCTC11168 variant 2 CPS. The structure of the CPS is shown above the spectra. (a) ^1H spectrum of the purified CPS. (b) Selective TOCSY (H-3C, 50 Hz, 80 ms) for assignments of the proton resonances of residue C. (c) Selective NOESY (H-4C, 50 Hz, 200 ms) to detect inter-residue NOEs between residue C and D. (d) Trace from the ^1H - ^{31}P HMQC for the ^{31}P signal at 13.6 ppm. (e) ^1H - ^{13}C HMQC spectrum showing assignments for residue C, the anomeric resonances and the POME resonance.

FIGURE 4 Is a depiction of MS analysis for the *C. jejuni* NCTC11168 variant 2 CPS. (a) CE-MS (m/z 100-1600) with orifice voltage of 200 V. (b) MS/MS spectrum of m/z 884 prompted by front-end collision induced dissociation.

- 5 **FIGURE 5** Is a depiction of strategy for amplification of the *cps* regions. The primers corresponding to the conserved *cps* genes (shown in open arrows) were used in combination with primers derived from conserved regions of *kps* genes (thick solid arrows) for long-range PCR as described in the section Experimental Procedures.
- 10 **FIGURE 6** Is a depiction of graphical representation of the sequenced CPS biosynthethic regions. In cases with high level of similarity between putative gene products (usually with e-values below 1e-30), the genes were given names of counterparts found in other bacteria. When no such similarity was found, the genes were assigned the names of respective genes from strain NCTC11168. The genes with
- 15 no similarity to either NCTC11168 or other bacteria are given strain-specific systematic names. The *cps* clusters of serostrains HS:23 and HS:36 are almost identical to that of strain 81-176 and are not shown (see text).

- FIGURE 7** Is a depiction of summary of the *C. jejuni* capsular polysaccharide
- 20 structures described in this study. The CPS structures of the heat-stable (HS) Penner type strains HS:1, HS:19, HS:23 and HS:36 have been reviewed by Moran *et al.* {Moran, 2000 #217}. The structures of NCTC11168 CPS {St Michael, 2002 #212} and HS:41 CPS {Hanniffy, 1999 #175} have recently been described. Sugars are shown in pyranose configurations unless otherwise noted. P, phosphate; Gal,
- 25 galactose; Gro, glycerol; Me, methyl; Hep, heptose; Rib, ribose; GalNAc, N-acetylgalactosamine; GlcA6, glucuronic acid; NGro, aminoglycerol; GlcNAc, N-acetylglucosamine; Ara, arabinose; Alt, altrose; Fuc, fucose.

- FIGURE 8** Is a depiction of proton NMR spectra of *C. jejuni* strains NCTC12500
- 30 (HS:1 serostrain) and G1 (HS:1). a) HR-MAS spectrum of NCTC12500 (HS:1) whole cells at 21°C. b) HR-MAS spectrum of G1 whole cells at 21°C. c) NMR spectrum of partially purified G1 CPS at 40°C with acetone as the internal reference. The CPS

anomeric resonance corresponding to Gal and the methyl resonance from the common phosphoramidate are labeled.

FIGURE 9 Is a depiction of proton NMR spectra of *C. jejuni* strains CCUG 10954 (HS:23 serostrain), ATCC 43456 (HS:36 serostrain) and 81-176 (HS:23/HS:36). a) NMR spectrum of partially purified CCUG 10954 CPS. b) NMR spectrum of ATCC 43456. c) NMR spectrum of 81-176. HR-MAS spectra were acquired at 21°C. NMR spectra were acquired at 40°C. For the CPS, acetone was used as the internal reference. **Is also a depiction of the HMQC and TOCSY spectra of the** *C. jejuni* strains CCUG 10954 (a), ATCC 43456 (b), and 81-176 (c). For the HMQC spectra, the anomeric region is shown. For the TOCSY, the mixing time was 90ms. Crosspeaks between signals in the anomeric region (4.7 to 5.5 ppm) and the sugar ring region (3.4 to 4.4 ppm) are shown.

FIGURE 10 Is a depiction of comparisons of the *cps* clusters of 81-176 (top) with NCTC11168 (middle) and NCTC12517 (bottom). Each of the genes are shown as boxes.

FIGURE 11 Is a depiction of regions of significant homology between the capsule locus (Cj1413c to Cj1448c) in *C. jejuni* NCTC11168 and *C. jejuni* RM1221, *Helicobacter hepaticus* ATCC51449, *Helicobacter mustelae* ATCC43772, and *Wolinella succinogenes* DMSZ 1740. No significant homology was seen with *Helicobacter pylori* strains 26695 or J99. Note that the genome sequences of *C. jejuni* RM1221 and *H. mustelae* ATCC43772 are incomplete.

FIGURE 12 31P HMQC NMR analysis of purified capsule isolated from *C. jejuni* HS:1. The signal at 3.81 ppm (arrow) is characteristic of a phosphoramidate modification found on the capsule and originates from the CH₃ methyl group of the phosphoramidate. The signal at 4.84 ppm (arrow) is indicative of the capsular sugar to which the phosphoramidate is attached. This sugar was not described in previous structural studies of HS:1 and is currently under investigation.

FIGURE 13 ^{31}P HMQC NMR analysis of purified capsule isolated from *C. jejuni* HS:19. The signal at 3.76 ppm (arrow) is characteristic of a phosphoramidate modification found on the capsule and originates from the CH_3 methyl group of the phosphoramidate. The signal at 4.26 ppm (arrow) is indicative of the capsular sugar to which the phosphoramidate is attached (position 4 of β -GlcNAc).

FIGURE 14 Whole cell ^{31}P - ^1H filtered ^1H HR-MAS NMR spectra of several strains of *C. jejuni*. The number of peaks represent the number of phosphoramidate residues in different chemical(structural) environments. The amplitude of the peaks reflects the relative amounts of each residue.

FIGURE 15 Adherence and invasion results for phosphoramidate mutants *C. jejuni* 1416-1, 1417-1, 1418-3 and their parent strain, NCTC11168H (UK-H) as well as the phosphoramidate mutant *C. jejuni* 1421-3 and its parent strain, Variant 4 (V4). Efficiency is defined as the number of bacteria that either adhered or invaded CaCo-2 cells divided by the total number of bacteria added and expressed as a percentage. Results are presented as the means of at least three experiments $\pm$ the standard error of the mean.

FIGURE 16 Results of motility assays for phosphoramidate mutants *C. jejuni* 1416-1, 1417-1, 1418-3 and their parent strain, NCTC11168H (UK-H) as well as the phosphoramidate mutant *C. jejuni* 1421-3 and its parent strain, Variant 4 (V4). Results are presented as the mean of at least two experiments $\pm$ the standard deviations.

FIGURE 17 Results of two serum sensitivity assays for the phosphoramidate mutant *C. jejuni* 1416-1 and the parent strain, NCTC11168 (UK-H). Bars in grey represent colony counts without serum while bars in white represent counts in the presence of 100 μl of serum.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

While the invention is discussed with respect to particular examples and embodiments, it will be readily understood that it is not so limited, but in fact includes all variants and alternative embodiments thereof. While possible mechanisms and/or

modes of action may be discussed, it will be understood that the invention is not so limited.

- 5 The inclusion of a reference is not an admission or suggestion that it is relevant to the patentability of anything disclosed herein.

Bacterial strains and growth conditions - *Campylobacter jejuni* NCTC11168 (HS:2) was isolated from a case of human enteritis and later sequenced by Parkhill *et al.* *C. jejuni* serostrains: HS:1 (ATCC 43429), HS:2 (ATCC 43430), HS:3 (ATCC 43431), HS:4 (ATCC 43432), HS:10 (ATCC 43438), HS:19 (ATCC 43446), HS:36 (ATCC 43456) and HS:41 (ATCC 43460) were obtained from ATCC; *C. jejuni* HS:23 was obtained from Dr. Peggy Godschalk, Erasmus University Medical Center, Rotterdam; *C. jejuni* OH4382 and OH4384 were obtained from Health Canada; and
10 *C. coli* HS:30 (NCTC 12532) was obtained from NCTC. All campylobacter strains were routinely grown on Mueller Hinton agar (Difco) under microaerophilic conditions at 37°C. *C. jejuni* NCTC11168 mutants were grown on Mueller Hinton agar with 30 µg/mL kanamycin.

Spectroscopy – All CE-ESI-MS and CE-ESI-MS/MS experiments and structural analysis of the purified CPS by NMR were performed substantially as described (in St. Michael *et al.* Eur.J. Biochem 269:5119, (2002)). ³¹P NMR experiments were acquired using a Varian Inova 500 MHz spectrometer equipped with a Z-gradient 3 mm triple resonance (¹H, ¹³C, ³¹P) probe substantially as described in Kneidinger *et al.* JBC 278, 3615 (2003). External 85% phosphoric acid
20 was used at the chemical shift reference.

Preparation of cells for HR-MAS NMR - *C. jejuni* overnight growth from one agar plate (~10¹⁰ cells) was harvested and suspended in 1 mL of 10 mM potassium buffered saline (pH 7) made in D₂O containing 10% sodium azide (w/v). The suspension was incubated for 1 h at room temperature to kill the bacteria. The cells
30 were pelleted by centrifugation (7 500 X g for 2 min) and washed once with 10 mM potassium buffered saline in D₂O. The pellet was resuspended by adding 20 µL of D₂O and then 40 µL of the suspension was inserted into the rotor for analysis.

HR-MAS NMR spectroscopy – HR-MAS experiments were performed using a
35 Varian Inova 600 MHz spectrometer equipped with a Varian nano-NMR probe

substantially as described in St. Michael (2002), above, and Young *et.al.* JBC 277:42530 (2002). Spectra from 40 μ L samples were spun at 3 KHz and recorded at ambient temperature (21°C). The experiments were performed with suppression of the HOD signal at 4.8 ppm. Proton spectra of bacterial cells were acquired with the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence $[90-(\tau-180-\tau)_n\text{-acquisition}]$ to remove broad lines arising from lipids and solid-like material. The total duration of the CPMG pulse ($n*2\tau$) was 10 ms with τ set to (1/MAS spin rate). One-dimensional selective TOCSY experiments with various spin-lock times from 30-150 ms and selective NOESY with mixing times from 100-400 ms were performed substantially as described in Uhrin and Brisson (2000) in NMR in Microbiology, p.165-210 Horizon Science Press, UK, and in Brisson *et.al.* (2002) in NMR spectroscopy of glycoconjugates p.59-93, Wiley-BCH, Weinheim. For use under MAS conditions, the TOCSY sequences were modified so that the DIPSI-2 mixing sequence was replaced with the adiabatic WURST-2 pulses. Selective experiments were described as EXP[selected spins, selective excitation bandwidth, mixing time] where EXP is TOCSY or NOESY. Typically, proton spectra of bacterial cells could be obtained using 256 to 1024 transients (15 min to 1 hour). For the selective experiments on the N-linked glycan resonances present as a minor component in the bacterial cells, the time for each TOCSY and NOESY varied from 1 to 8 hours.

Deoxycholate-PAGE, silver-staining and immunoblotting of polysaccharides - Proteinase K treated whole cells of *C. jejuni* wild type and phase variants were prepared and analyzed by deoxycholate-PAGE substantially as described in St Michael (2002) above. One portion of the gel was silver-stained while the other portion of the gel was transferred to a PVDF membrane (Roche Molecular Biochemicals) and immunodetected with HS:2 antiserum (1:500 dilution). The immunoblot was then incubated with goat-anti-rabbit secondary antibody conjugated to alkaline phosphatase (1:2500 dilution, Sigma) and then developed with the nitro blue tetrazolium chloride / 5-bromo-4-chloro-3-indolyl phosphate detection system (Roche Molecular Biochemicals).

Bacterial strains and growth conditions - Sequences of *cps* regions from a wide range of *C. jejuni* strains have been investigated: NCTC11168 (HS:2, genome sequenced strain, enteritis isolate), 176.83 (HS:41 serostrain, enteritis isolate, NCTC12517 (HS:19 serostrain, enteritis isolate), G1 (HS:1, GBS isolate), 81-176

(HS:23/36, enteritis isolate used in human challenge studies), CCUG 10954 (HS:23 serostrain, enteritis isolate) and ATCC 43456 (HS:36 serostrain, enteritis isolate). Three of the strains examined in this study (NCTC11168, G1 and serostrain HS:19) and the additional strains used for comparative analysis of homopolymeric tracts in the *cps* region are listed in Table IV. *C. jejuni* strains were grown in microaerophilic conditions at 37°C on 7% blood agar plates for 2 days. The *E. coli* XL2 Blue MRF' strain (Stratagene), used in cloning experiments, was grown overnight at 37°C on LB agar plates supplemented with 100 ug/ml ampicillin when necessary.

Sequencing of homopolymeric tracts in contingency genes - Genes *cj1420* and *cj1421* were amplified with primers ak149 (GAGTGCCACTGCTTACACGAGC SEQ. ID. NO.1) and ak150 (GCTCAACCCAAATTCAGCCATAGAAAG SEQ. ID. NO.2) and sequenced with primers ak152 (CACCTCCTTTATACCAATTCTGATAAGCC SEQ. ID. NO.3) and ak151 (GGCATAAAGGGAGTGGCGAAGAAACCTGC SEQ. ID. NO.4), respectively. Gene *cj1422* was amplified using primers ak153 (GATACGGCACAGTAAATGTTGATG SEQ. ID. NO.5) and ak147 (GCTACTATATCTGGACGATGTTGCTTG SEQ. ID. NO.6) and sequenced with primers ak150 and ak152. Genes *cj1426* and *cj1429* were amplified using primers ak144 (CTCATTCGACCTTTGGAATTGCCTTTG SEQ. ID. NO.7) and ak145 (CTGTTTCATAATTTCTGTCCGATACTGC SEQ. ID. NO.8) and sequenced using the same primers. Gene *cj1437* was amplified with primers ak154 (CTCCTTATTTATCTATTCCACAC SEQ. ID. NO.9) and ak155 (CTTGTATTTTTTCAGCAACATAACTC SEQ. ID. NO.10) and sequenced with primer ak156 (GCATTGGCGAGTTTTAGGATAGG SEQ. ID. NO.11). The single base run polymorphisms are detected directly from the sequencing chromatograms, from which the variable number Gs for a particular region can be estimated via comparison with the published consensus sequence.

The strategy of amplification and sequencing of C. jejuni CPS regions - In the preliminary experiments using PCR analyses, high sequence conservation of the CPS transport and assembly genes (*kps* genes,) was found flanking the internal biosynthetic region of the CPS locus (*cps* cluster). In some cases such conservation extended into the adjacent internal *cps* genes. As all CPS-related genes in strain NCTC11168 are transcribed in the same direction, the strategy of PCR amplification

was based on the assumption that this is the case in all other *C. jejuni* strains. In order to design primers suitable for long-range PCR, short sequences of *kpsC* and *kpsF* genes using the SP-PCR procedure were derived. Sequence comparison allowed detection of highly conserved regions suitable for the design of universal PCR
5 primers.

Primers ak176 (CGGTTACCGCTTAACACATCAGGATGGGG SEQ. ID. NO. 12) and ak177 (GTAAACCCCAGCCCGCATAAAAAGGC SEQ. ID. NO.13) were used in SP-PCR sequencing of *kpsC* genes, and primers ak173 (GGGCGTTGCATAGTTAGTGGTATGGGTAAATCAGG SEQ. ID. NO.14) and
10 ak174 (GGCGCTAAGATAGCAGCTACTTTAGCAAGCACAGG SEQ. ID. NO. 15) were used for SP-PCR sequencing of fragments of *kpsF* genes in various strains of *C. jejuni*. Alignment of the derived sequences allowed the design of universal primers suitable for long PCR: ak186 for *kpsF* gene and ak188 for *kpsC* (see below). For serostrain HS:19 a more optimal *kpsC* primer ak187 was designed.

15 Long-range PCR with *kpsF* and *kpsC* primers alone failed to produce any product with the reference strain NCTC11168. This could be due to a relatively large size of the amplicon (over 36 kb). However, it was possible to amplify the entire *cps* region as two long-range PCR products, when *kpsC* and *kpsF* primers were combined with primers derived from internal biosynthetic genes.

20 The same strategy was used for amplification of *cps* regions from other strains. A possibility of extension of the strategy to other strains was based on an assumption that various strains would share some genes in the internal *cps* clusters. This was supported by preliminary hybridisation data. It was assumed that the conserved genes, when present, would be located in the same orientation. The
25 sequencing of the *cps* regions of serostrains HS:23 and HS:36 was performed after the sequencing of strain 81-176 was complete. The identical gene content and high sequence identity between these three strains allowed complete sequencing of the *cps* regions of serostrains HS:23 and HS:36 using custom-made oligonucleotides.

The biosynthetic *cps* region of strain NCTC11168 contains 28 genes. The
30 conserved genes present in the biosynthetic regions of other strains were identified using PCR amplification with primers derived from the sequence of NCTC11168 genome. When a product of expected size was present, the corresponding primers could be used in combination with the *kpsC* and *kpsF* primers for long-range PCR.

Primer pairs specific to the biosynthetic *cps* genes of strain NCTC11168 are indicated in the supplementary material. The primers designed for the genes found in internal *cps* regions of various strains were used in combination with the primers corresponding to the conserved flanking *kpsC* and *kpsF* genes to generate long PCR products (Fig. 5). The long-range PCR resulted in overlapping products suitable for generation of complete sequences of the internal biosynthetic regions. Long-range PCR was performed using the Expand 20 kb^{PLUS} PCR System (Roche) using conditions described by the manufacturer. *KpsC* specific primers ak188 (CCCCTAAAATCATCGAAGCATCATCTTCAACTTGAGC SEQ. ID. NO.16) and ak187 (CATGCTTTAAACCATTATACTTTGAAAAGCGGTTCTCAAG SEQ. ID. NO.17), for serostrain HS:19) and *kpsF* specific primer ak186 (GAAAAGGAAGCTTGTCTTTGCAGCTTGC SEQ. ID. NO.18) were used in long-range PCR experiments.

The long-range PCR products were treated with polynucleotide kinase, sonicated, and blunt-ended with T4 DNA polymerase. Then, 1-2 kb fragments were gel extracted and cloned into alkaline phosphatase-treated pUC18 (Promega) prior to sequencing. For closing gaps, primers corresponding to the ends of the contigs were designed and the regions were amplified and sequenced either directly or after cloning into pGEM-T-Easy vector (Promega) using the automatic sequencer. DNA sequencing was performed on ABI 377 or ABI 3700 automatic sequencers using an ABI PRISM BigDye Terminator Cycle Sequencing Kit (Perkin-Elmer). The sequences generated via shot-gun sequencing were assembled and edited using GAP4 or GeneTool software (DoubleTwist.com), and were deposited at EMBL database with the following IDs and accession numbers:

Designation	Strain	ID	Acc. Number
HS19	NCTC12517	CJ12517CPS	BX545860
HS1	G1	CJG1CPS	BX545859
HS23/36	81-176	CJ81176CPS	BX545858
HS41	176.83	CJ17683CPS	BX545857
HS:36	ATCC 43456	CPS_O36	AY332624
HS:23	CCUG 10954	CPS_O23	AY332625

Multiple sequence alignment was performed using the ClustalW program (<http://www2.ebi.ac.uk/clustalw/>). The *cps* sequences were analysed using Artemis software and the extracted amino acid sequences were analysed by similarity searches with the BLASTp program against NCTC11168 at http://www.sanger.ac.uk/Projects/C_jejuni/ and a non-redundant protein database at <http://www.blast.genome.ad.jp/>. The entire *cps* regions were compared with the *cps* region of NCTC11168 using BLASTn and tBLASTx programs (<http://www.hgmp.mrc.ac.uk/>) followed by the analysis using MSPcrunch (<http://bioweb.pasteur.fr/seqanal/interfaces/mspcrunch.html>) and ACT programs (<http://www.sanger.ac.uk>).

Isolation and purification of CPS - The CPS was isolated from dried cell mass (approx. 1 g) by the hot water/phenol method {Westphal, 1965}. The aqueous phase was dialyzed against water and lyophilized. The dried sample was then dissolved in water to a 1 % solution (w/v) and subjected to ultracentrifugation to yield a gel-like pellet containing LOS and supernatant containing the CPS.

Analytical methods - Sugars were determined by examining their alditol acetate derivatives by GLC-MS. Samples were hydrolyzed for 4 h using 4 M trifluoroacetic acid at 100°C. The sample was reduced in NaBD₄ overnight in H₂O and acetylated with acetic anhydride at 100°C for 2 h using residual sodium acetate as the catalyst. The GLC-MS was equipped with a 30 M DB-17 capillary column (180°C to 260°C at 3.5°C/min) and MS was performed in the electron impact mode on a Varian Saturn II mass spectrometer.

HR-MAS NMR allows the screening of small amounts of bacterial cells directly without having to purify surface carbohydrates. HR-MAS experiments were performed on a Varian Inova 600 MHz spectrometer using a gradient 4 mm indirect detection high-resolution magic angle spinning nano-NMR probe (Varian) with a broadband decoupling coil as previously described {St Michael, 2002; Young, 2002}. Proton spectra of cells were acquired with the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence {90-(τ -180- τ)_n-acquisition} to remove broad lines arising from lipids and solid-like material. The total duration of the CPMG pulse ($n2\tau$) was 10 ms with τ set to (1/spin rate).

High-resolution NMR experiments on the partially purified CPS were acquired using a Varian Inova 500 MHz spectrometer equipped with a Z-gradient 3

mm triple resonance (^1H , ^{13}C , ^{31}P) probe. The experiments were performed at 40°C with suppression of the water resonance. The methyl resonance of acetone was used as an internal reference at δ_{H} 2.225 ppm and δ_{C} 31.07 ppm. Standard sequences from Varian, COSY, TOCY, NOESY, HMQC, and ^{31}P HMQC were used.

5 *Adherence and invasion assays* - The procedure was performed as described previously by Bacon *et al.* (2001) with the following modifications: there were approximately 2×10^5 CaCo-2 epithelial cells per well infected with approximately 3×10^7 bacteria (multiplicity of infection = 150 bacteria per epithelial cell).

10 *Motility assays* - The OD_{600} of a suspension of bacterial cells in Mueller-Hinton broth was adjusted to 1.0 and $5\mu\text{L}$ of the culture was inoculated into the centre of duplicate 0.4% agar plates. The plates were incubated at 37°C under microaerophilic conditions and the diameter was measured after 52 hours.

15 *Serum sensitivity assays* - Five μL of a bacterial suspension, adjusted to an $\text{OD}_{600} = 0.1$, is added to duplicate wells containing $900\mu\text{L}$ of Mueller-Hinton (MH) broth with either $100\mu\text{L}$ of active serum or additional MH broth. After a 1-hour incubation under microaerophilic conditions at 37°C with shaking at 100rpm, strains from each well are diluted to 10^{-2} and 10^{-3} and plated on dry MH agar plates. The plates are incubated under microaerophilic conditions for 2 days and colony counts are performed.

20

RESULTS

25 *Examination of CPS from whole cells by HR-MAS NMR* - The capsular polysaccharide structure of the genome sequenced strain, NCTC11168 (HS:2) was described. The proton spectrum obtained from HR-MAS of suspended NCTC11168 bacterial cells closely resembled the spectrum of the purified capsular polysaccharide and clearly demonstrated the *N*-acetyl, *O*-methyl, and anomeric resonances (Fig. 1a, b). The HR-MAS NMR spectrum was obtained in a few minutes directly from 40 μL of whole cells. Hence, this method permitted quick screening of campylobacter CPS directly from one plate of growth ($\sim 10^{10}$ cells), but was sensitive enough to detect a
30 1/100 dilution of the suspension containing 8×10^7 cells (Fig. 1c).

This method also permitted serotype comparisons between strains. The capsular polysaccharide structure from the HS:2 serostrain had not been determined.

In fact, it was previously believed that this strain did not produce high molecular weight glycans. Whole cell NMR spectra of the HS:2 serostrain and NCTC11168 are comparable (Fig. 1a,d). These results provide further evidence that capsular polysaccharides are the main serodeterminant in the heat-stable typing scheme and demonstrate that HR-MAS NMR can be used to confirm serotype.

While simple HR-MAS spectra of bacterial cells can allow one to monitor glycan resonances, assignment of resonances to specific residues may require further information. In the present study, selective TOCSY and NOESY experiments were employed to identify sugar residues or assign unknown resonances. For example, in the spectra of several campylobacter strains, sharp multiplets were often observed between 2.6 and 2.9 ppm (Fig. 1d). A series of selective TOCSY experiments identified all the spins for this compound and they were determined to correspond to those of free aspartic acid. Addition of aspartic acid to the cells resulted in increased peak intensity between 2.6 and 2.9 ppm.

Also, selective TOCSY or NOESY could be performed on various *C. jejuni* serostrains to identify other sugar resonances in accordance with those reported in the literature. In the case of *C. jejuni* HS:41, previous studies have reported that the purified CPS was composed of a mixture of polysaccharides with the major and minor components differentiated by the presence of either 6-deoxy-altrofuranosyl or D-fucofuranosyl residues. The HR-MAS spectrum of HS:41 cells exhibited extensive spectral overlap so that signals for the signature 6-deoxy-sugars could not be assigned unambiguously. However, selective TOCSY experiments on HS:41 cells starting with selective irradiation of ^1H resonances near 1.2 ppm established scalar coupling connectivities between the dominant CH_3 resonance at 1.27 ppm with other sugar ring protons whose chemical shifts (5.23 ppm, 4.25 ppm, 3.9 ppm, 3.7 ppm) agreed with those of the 6-deoxy-D-altrofuranosyl moiety reported for the purified CPS (5.185 ppm, 4.19 ppm, 4.34 ppm, 3.71 ppm, 3.89 ppm, and CH_3 1.27 ppm). These results suggests that the dominant form of the HS:41 CPS contains the 6-deoxy-altrofuranosyl moiety.

HR-MAS NMR analysis of C. jejuni NCTC11168 CPS phase variants –10 single colonies of NCTC11168 were selected and restreaked to one plate each. The growth from a single plate was examined directly by HR-MAS NMR and also digested with proteinase K followed by deoxycholate-PAGE silver-staining or

immunoblotting. Three different phenotypes were observed by these methods. Growth from the first colony (variant 1) showed similar silver-staining patterns relative to the wild type population from which it was isolated from but showed increased levels of reactivity with HS:2 sera (Fig. 2a, b). Comparison of the HR-MAS NMR spectra (Fig. 2) with those of the wild type NMR spectra of purified CPS demonstrated that variant 1 predominantly exhibited a resonance at 3.2 ppm consistent with an *N*-ethanolamine modification on the glucuronic acid (GlcA) (Fig. 2d) in contrast to the major wild type form which exhibited GlcA modified with aminoglycerol (Fig. 2c). Variant 2 showed extremely reduced levels of silver-staining and immunoblotting although the HR-MAS spectra clearly indicated that similar amounts of polysaccharides were present in both variant and wild type samples (Fig. 2e). The HR-MAS spectrum of variant 2 revealed new resonances at 3.75 ppm (Fig. 2e) indicative of a novel modification, which had not been previously observed. In addition, the anomeric chemical shift for residue C moved downfield closer to the one for residue B. Variant 3 showed increased silver-staining but similar levels of immunoreactivity. This variant lacks the 6-O-Me group on the heptose confirmed by the loss of the resonance at 3.55 ppm (Fig. 2f).

The structural determination of the purified polysaccharide from *C. jejuni* NCTC11168 variant 2 was done substantially as described in St. Michael (2002), above. Its backbone CPS structure was found to be the same as determined previously (Fig. 1), but with the addition of a modified phosphate group at C-3 of the Gal/NAc residue C (Fig. 3). The proton spectrum of the purified CPS from variant 2 is shown in Fig. 3a. The sample also contained about 30% of the major wild-type CPS whose structure is shown in Fig. 1. Comparison of the HMQC spectra of variant 2 with the one from the wild-type sample, showed similarity in chemical shifts for residues A, B and D (Fig. 3e and Table I). Proton chemical shifts for residue C for variant 2 were identified using a selective TOCSY experiment (Fig. 3b). The 5D-4C and 3D-4C NOEs were also observed (Fig. 3c), as before for the wild-type CPS.

The proton spectrum for variant 2 (Fig. 3a) contained a signal for the methyl group linked to phosphate via an ester bond: δ_H 3.75, δ_C 54.8 ppm, which had a $J_{P,H}$ of 11.5 Hz. In the ^{31}P - 1H HMQC spectrum (Fig. 3d) this methyl group showed correlation to the ^{31}P signal at 13.6 ppm, which also gave a correlation to H-3 of the Gal/NAc residue C. A $J_{P,H-3C}$ value of 8 Hz was obtained from simulation of the

undecoupled ^{31}P spectrum. In the ^{31}P - ^1H HMQC-TOCSY spectrum, correlations from the phosphorus signal at 13.6 ppm to H-2, H-3, H-4, and H-5 of the Gal/NAc residue were observed. These data indicated that the methylphosphate group was linked at O-3 of the Gal/NAc residue C. However, the low field chemical shift of ^{31}P resonating at 13.6 ppm was inconsistent with the presence of a phosphodiester group.

Different conditions of mild acid hydrolysis of the polysaccharide were tested, in an effort to cleave furanoside bonds without complete destruction of the phosphor-containing substituent. Hydrolysis with 1% trifluoroacetic acid for 20 min. at 100° C completely depolymerized the polysaccharide. At the same time the ^{31}P signal at 13.6 ppm disappeared and a group of ^{31}P signals with one major component arose at ~2 ppm. All of them correlated with methyl group signals, with the major methyl signal observed at 3.88 (^1H)/56.7 (^{13}C) ppm. Milder hydrolysis conditions (1% TFA at 60° or 2% AcOH at 100°, 1h) led to incomplete conversion of the phosphate group, but did not depolymerize the polysaccharide completely. However, the chemical behavior and ^{31}P chemical shift of the methylphosphate group were consistent with the presence of the amide of methylated phosphoric acid $\text{R-OP=O}(\text{NH}_2)\text{OMe}$. Phosphoramides can be hydrolyzed in dilute acids with the replacement of the NH_2 group with the OH group, in the conditions where alkyl esters of phosphoric acid are stable. Phosphoramides usually have ^{31}P signals between 10 and 20 ppm (41-47), which agrees with the position of the ^{31}P signal within the analyzed structure.

CE-MS analysis of C. jejuni NCTC11168 CPS phase variant 2 – In order to confirm the structure of the capsular glycan derived from NMR studies, the purified CPS sample was also analyzed by using CE-MS and CE-MS/MS techniques. All the CE-MS and CE-MS/MS experiments were acquired using high orifice voltage. With this experimental setup, the polysaccharide breaks up into shorter oligosaccharide units due to the front-end collision induced dissociation. In this study, a orifice voltage of 200 V was applied and the extracted mass spectrum is shown in Fig. 4a. Compared to the spectrum obtained with a low orifice voltage (60 V), the typical polymer peak disappeared and strong peaks that correspond to oligosaccharide or monosaccharide units appear. The ion $[\text{M} + 1\text{H}]^{1+} = 884$ corresponded to the mass of the one repeat unit minus H_2O . The ion m/z 791 arose from the repeat unit of the wild-type polysaccharide lacking unit E (30% of the sample) or the loss of unit E.

The ions m/z 1181 and 1472 were assigned to one repeat unit plus CE, and to one repeat unit plus BD and A, respectively.

To further investigate the composition of the CPS repeat unit, the MS/MS experiments were conducted with the precursor ions at m/z 297, 382, 678, 884, 1181 and 1472. The MS/MS of ion m/z 297 clearly indicated the composition of C and E, whereas the MS/MS of ion m/z 382 displayed the composition of A and B. Although the presence of the sugar residues D and A were not directly detected (Fig. 4a), the existence of these two residues in the oligosaccharide repeat unit from the tandem mass spectrum of ion m/z 884 could easily be determined. As shown Fig. 9b, there was a lost of 206 Da, which corresponded to the mass of residue D, resulting in the fragment ion m/z 678. Similarly, the fragment ions m/z 588, 456 and 429 were generated from the losses of CE (296 Da), AC (428 Da), and BD (455 Da), respectively. In Fig. 4b, the fragment ion at m/z 186 was assigned to the anhydrate C (203 Da) and the ions at m/z 168 and 126 were generated from the consecutive neutral losses of H_2O (18 Da) and acetyl group (42 Da), respectively. The reason for the co-existence of fragment ions that correspond to losses of D, CE, ACE, and BD could be explained by the nature of generation of ions m/z 884 which is due to different breakage points along the polymer chain (A[CE][BD], [CE][BD]A, [BD]A[CE]). Hence, all the MS and MS/MS data was consistent with the structure for the CPS of variant 2 shown in Fig.3.

Correlation of cps genes from strains of serotypes HS:1, HS:2, HS:19, HS:23, HS:36 and HS:41 with respective CPS structures - The strategy used for sequencing the variable *cps* loci from the different strains is described elsewhere herein and shown in Fig. 5. The overall summary of the *cps* sequencing results is presented in Table II. A schematic of all the *cps* loci compared in this study is shown in Fig. 6 with the genes involved in phosphoramidate biosynthesis shown in bold (see below). Some of the gene products are involved in the biosynthesis of activated sugars. Such activated sugars contain energy-rich nucleotide-phosphate bonds and serve as substrates for glycosyltransferases involved in the biosynthesis of polysaccharides. In addition, nucleotide sugars may be modified by enzymes such as epimerases, dehydratases and reductases before transfer of the final product. Additional modifying enzymes can add groups such as O-methyl, phosphate, ethanolamine- and aminoglycerol- to further increase the complexity of the structures. Indeed, genes

encoding these enzymes can be found in various *C. jejuni* *cps* regions. The predicted function of *cps* genes from strain NCTC11168 (HS:2) based on the published genome sequence {Parkhill, 2000} and the recently published CPS structure {St Michael, 2002} are presented in Table III. The CPS structures of NCTC11168 and the other strains used in this study are shown in Fig.7.

There are three notable features encoded by the *cps* locus of NCTC11168: homologues of the GDP-D-*glycero*-D-*manno*heptose pathway (HddC, GmhA2 and HddA), the presence of a UDP-glucose dehydrogenase homologue, Udg, responsible for the formation of UDP-glucuronic acid, and a UDP-pyranose mutase homologue, Glf, catalysing the reversible conversion of pyranoses to furanoses.

NMR analysis of the HS:19 serostrain used in this study confirmed that the CPS structure was consistent with the published disaccharide repeat (Fig.7, results not shown). The *cps* region of the HS:19 serostrain did not contain homologues of the heptose pathway but did have the *udg* homologue (Table IV) correlating well with the presence of β -D-glucuronic acid which is also amidated with 2-amino-2-deoxyglycerol. NMR analysis also detected two acid-labile functional groups that were not reported previously. Both the phosphoramidate modification recently described for NCTC11168 and an unknown labile group were observed during the analysis.

In contrast to NCTC11168 (HS:2) and the HS:19 serostrain, the CPS locus of G1 (HS:1) does not encode a homologue of UDP-glucose 6-dehydrogenase (Table V) and thus the strain should not have the ability to synthesise glucuronic acid. However, this strain contains a potential *tagD* homologue encoding a glycerol-3-phosphate cytidyltransferase necessary for the formation of CDP-glycerol (Table V, G1.11). G1 also encodes a TagF homologue, which transfers glycerol-phosphate residues from CDP-glycerol. Therefore, the repeating unit of this CPS may contain glycerophosphate residues. Indeed, the HS:1 serostrain was reported to contain glycerol-1-phosphate residues alternating with galactose in the repeating unit (Fig. 7). The NMR spectra of G1 (Fig. 8) revealed that the structure of this CPS is consistent with the HS:1 structure. An additional anomeric resonance in the HR-MAS spectrum of G1 was not present in the partially purified CPS sample suggesting that this resonance probably came from the medium used. Extensive NMR analysis by COSY, TOCSY, NOESY and HMQC indicated the presence of only one anomeric resonance

consistent with the presence of one sugar in the repeating unit. ^{31}P NMR experiments indicated the presence of a phosphate diester linkage, also consistent with the reported structure. The common anomeric resonances corresponding to the *N*-linked Pgl glycan were also observed during HR-MAS analysis for both the HS:1 serostrain and
5 G1, although they are prominent only for the HS:1 serostrain in Fig. 8. The phosphoramidate modification was also observed during analysis of both strains (Fig. 8). In the ^{31}P HMQC spectra, a strong correlation was observed between the POME resonance at 3.8 ppm and the phosphoramidate resonance at 14 ppm, indicative of a –
OP=O(NH₂)OMe modification.

10 The CPS loci of the HS:23 and HS:36 serostrains and of strain 81-176 (which reacts with both HS:23 and HS:36 antisera) all have exactly the same gene content (Figure 6 and Table II). The CPSs of HS:23 and HS:36 were found to contain repeating units of α -D-galactose, β -D-GlcNAc- and *D-glycero-D-altro*-heptose or deoxy variants with and without methyl groups (Fig. 9). However, it was reported that
15 the *D-glycero-D-altro*-heptose variant was not detected in the HS:23 serostrain. Analysis of the gene products, encoded by the *cps* regions of serostrains HS:23 and HS:36 and of strain 81-176 (Table VI), demonstrate a potential for deoxyheptose biosynthesis due to the presence of genes *hddC*, *gmhA2*, *hddA* and *dmhA* (Fig. 6). The latter gene homologue is suggested to be involved in conversion of heptose to
20 deoxyheptose in *Yersinia pseudotuberculosis*.

HR-MAS spectra of cells and NMR spectra of the partially purified CPS from strain 81-176 demonstrated similar sugar resonances with the HS:23 and HS:36 serostrains (Fig. 9). In all the spectra, the characteristic OMe signal at 3.5 ppm and NAc resonance at 2.05 ppm were observed. The anomeric region of the HS:23
25 serostrain was the simplest with two anomeric resonances in the HR-MAS spectrum (Fig. 9a). In the ^1H NMR spectra of the CPS, a third anomeric resonance was observed (Fig. 9b) which was obscured in the HR-MAS spectra by the large saturated HOD peak. The anomeric region (4.7 to 5.5 ppm) for the other samples was more complex with the spectrum of strain 81-176 being the most complex.

30 2D-NMR experiments were done to further characterize the sugar resonances (Fig. 9). The HMQC and TOCSY spectra for the HS:23 serostrain (Fig. 9a) were the simplest with proton anomeric resonances at 5.06 ppm, 4.97 ppm and 4.77 ppm, corresponding to the Gal, Hep and GlcNAc anomeric resonances, respectively. In the

HMQC spectrum, the C-6 crosspeaks of the 6-deoxy-heptose were observed at 34.8 ppm (^{13}C) and 2.06 and 1.71 ppm (^1H).

For the HS:36 serostrain, three anomeric resonances were also observed as detected by TOCSY and HMQC experiments on the CPS (Fig. 9b). The ^1H resonance
5 at 4.92 ppm was confirmed to be a non-anomeric resonance using HMQC. While the anomeric carbon resonances had similar chemical shifts, the proton anomeric resonance of the heptose residue was different, probably due to different structural motifs on the heptose residue. In the TOCSY spectrum, the anomeric resonances at 4.76 ppm and 5.06 ppm exhibited connectivities that were similar to those observed
10 for the HS:23 serostrain, indicating the presence of similar sugars in both serostrains. In the HMQC spectrum, resonances characteristic of a 6-deoxy-heptose could not be observed, indicating that for this serostrain this modification was not predominant.

The HMQC and TOCSY spectra for strain 81-176 (Fig. 9c) showed correlation patterns similar to those observed for the HS:23 serostrain for the Gal, Hep
15 and GlcNAc anomeric resonances, again indicating similar sugar structures to those of HS:23 and HS:36. This observation is in agreement with predictions derived from the 81-176 gene analysis. However, structural analysis of strain 81-176 also demonstrated the presence of additional resonances indicating the presence of a more complex repeating unit or the presence of another polysaccharide structure. Comparison of the
20 NOESY spectra established that serostrains HS:23 and HS:36 and strain 81-176 exhibited similar NOE patterns for the Gal, Hep and GlcNAc residues, a result that is consistent with the conclusions arrived at from the analysis of the TOCSY experiments. The phosphoramidate modification observed for NCTC11168 was also observed for serostrain HS:36 and strain 81-176 but not for serostrain HS:23.

25 The major and minor components of CPS isolated from the HS:41 serostrain were described to contain β -L-arabinose, 6-deoxy- β -D-altrioheptose, 6-deoxy- β -L-altriose and β -D-fucose all in the furanose form (Fig. 7). NMR analysis demonstrated that the CPS of the sequenced strain used in this study is consistent with the published structure. Interestingly, sequencing results from this strain (Table VII) show three
30 UDP-pyranose mutase *glf* gene homologues which are involved in pyranose to furanose ring conversions which is consistent with having three of the CPS sugars in the furanose form (note that arabinose is a pentose and therefore is naturally in the furanose configuration). The presence of genes: *hddC*, *gmhA2*, *hddA* and *dmhA*

(Fig.6) is consistent with the presence of deoxyheptose in the CPS (Fig.7). Additional sugar dehydratases will be required for the biosynthesis of fucose and deoxyaltrose and putative homologues are observed in Table VII. Two tandem copies of the *fcl* gene were also found in this strain. According to the CPS structure it appears that one copy is involved in heptose biosynthesis and the other in fucose production or alternatively, they are duplicate copies that are capable of converting both substrates.

Comparative analyses of the cps regions - The derived nucleotide sequences of various biosynthetic *cps* regions were compared with the complete genome sequence of strain NCTC11168 using both BLASTn and tBLASTx programs. Some features are outlined below.

The *cps* region of serostrain HS:19 contains genes which are almost identical to genes *cj1415-cj1420* of strain NCTC11168. This region (Fig.6) is followed by gene *HS19.07* with similarity in the 5' region to both gene *cj1421* and *cj1422*. However, no similarity between the 3' region of this gene and either the *cj1421* or *cj1422* genes of NCTC11168 could be found. Genes *cj1423-cj1433* are not present in serostrain HS:19 and there is limited similarity to genes *cj1434-cj1435*, *cj1437-cj1438* and *cj1440-cj1442* (Fig. 6 and Table IV).

The biosynthetic region of strain G1 is the smallest (15 kb) and contains only 11 genes. Organisation of the genes from *cj1415* to *cj1421* in this strain is similar to that of serostrain HS:19 and strain NCTC11168. However, the remaining genes have no counterparts in the corresponding regions of these strains (Fig.6, Table V).

Genes *cj1415-cj1420* are also conserved in the HS:23 serostrain, the HS:36 serostrain and strain 81-176 (Fig. 6, Table VI). However, in this case there are also a number of other conserved genes outside this region. Genes *cj1423* (*hddC*), *cj1424* (*gmhA2*), *cj1425* (*hddA*) and *cj1427* are conserved and present in the same place as in NCTC11168, but there is almost a precise deletion of gene *cj1426* (Fig. 6). Also, a new gene (*dmhA*) is present between genes *cj1427* and *fcl*. Genes *cj1429* and *cj1430* appear to be present (Fig. 6), but genes *cj1431* and *cj1440* appear replaced with two genes (*81176.16* and *81176.17* as well as the corresponding ORFs in the HS:23 and HS:36 serostrains) encoding glycosyltransferases. Overall, despite gene reshuffling, the *cps* regions of the HS:23 and HS:36 serostrains and strain 81-176 are more similar to that of NCTC11168 than to serostrain HS:19 and strain G1.

The *cps* region of serostrain HS:41 is interesting in that it lacks the *cj1415-cj1420* genes conserved in the other strains. However, three heptose-related genes in the middle of the *cps* locus of serostrain HS:41 (*hddC*, *gmhA2* and *hddA*) are almost identical to those in NCTC11168, although *gmhA2* and *hddA* are separated via
5 insertion of gene HS41.09 encoding a putative sugar transferase with low similarity to *cj1300* (Fig. 6, Table VII). The mosaic patterns of similarity and divergence indicate that these *cps* regions have a diverse recent ancestry, suggesting that recombination between different *cps* clusters has occurred.

The serostrains HS:23 and HS:36 and strain 81-176 (HS:23/36) all appeared to
10 have the same gene content. Pair-wise alignments of the CPS biosynthetic regions (24.6 kb) of these strains showed that the HS:23 and HS:36 serostrains share 97.6% DNA sequence identity between them while strain 81-176 shares 97.6% and 98.9% identity with the serostrains HS:23 and HS:36, respectively. As expected from the high DNA sequence identity in this region, there is also high protein sequence identity
15 when the individual ORFs are compared (see Table VI). All pair-wise comparisons showed above 93% protein sequence identity except for ORF *HS23.08* (*hddC*) which shared 87.9% and 86.6% identity with the corresponding ORFs in the HS:36 serostrain and strain 81-176, respectively.

Variation in the contingency genes - The potentially phase variable *cps* genes
20 of *C. jejuni* strains of various serotypes were investigated. The biosynthetic *cps* locus of *C. jejuni* NCTC11168 was found to contain six genes with homopolymeric G tracts potentially prone to phase variation. It was examined whether the “ON” and “OFF” states of these genes, if present in other strains, can also be detected. The results of this analysis, shown in Table VIII, indicate that most of the genes tested are
25 predominantly in the “ON” state although many are demonstrated to vary. This suggests that closer examination of these genes variable modifications (for example, methyl, ethanolamine, aminoglycerol and phosphoramidate) will reveal that CPS structures can be further modulated.

Such modulation may explain the presence of variant structures in the HS:41
30 serostrain and in serostrains HS:23 and HS:36 compared to strain 81-176 (HS:23/36). The latter three strains were examined in more detail because they share the same gene content, yet produce capsules with slight differences in CPS structure (Fig. 7). Since these three strains share >95% gene identity in their *cps* biosynthetic regions,

phase-variable genes could be responsible for the differential expression of deoxyheptose and phosphoramidate observed in this study, ie HS:23 (Gal, GlcNAc, Hep, **deoxyhep**), HS:36 (Gal, GlcNAc, Hep, **phosphoramidate**) and 81-176 (Gal, GlcNAc, Hep, **deoxyhep**, **phosphoramidate**). There are six contingency genes in the *cps* cluster of the HS:23 and HS:36 serostrains but only five in 81-176 since the *dmhA* homologue (ORF#12) is not phase-variable (Table IX). DmhA has been shown to be involved in deoxyheptose synthesis in *Yersinia* and interestingly, in this study, the *dmhA* homologue is functional in 81-176 and HS:23, but variable in HS:36. This may correspond to the detection of deoxyheptose in 81-176 and HS:23 and the difficulty in detecting this heptose variant in HS:36. In the HS:23 serostrain, two “OFF” genes (*HS:23.07* and *HS:23.20*) show high sequence similarity with the putative glycosyltransferase (*cj1422c*) from NCTC11168 and may play a role in adding the missing phosphoramidate. However, function of these contingency genes must be proven experimentally.

In this study, CE-MS/MS and HR-MAS NMR have been used successfully to examine glycan structures from 10^8 - 10^{10} bacterial cells. These methods can now be applied to investigate expression of glycans under different laboratory growth conditions and directly from the natural environments in which the pathogen is found. Examination of mutants will allow the assignment of genes involved in the biosynthetic pathways of these glycans and their modifications and help to determine the importance of structural phase variability in survival and pathogenesis. Due to the sensitivity and mildness of these methods, minor glycan structures that were not previously identified in the literature can also be detected.

Recently, HR-MAS NMR has been used to detect the polysaccharides of LPS and CPS on intact bacteria. The method has also been used to detect nanomole amounts of purified LPS, purified *O*-linked and *N*-linked glycopeptides, and LOS ganglioside mimics. In this study, HR-MAS NMR was used to further examine CPS directly from campylobacter cells. The CPS resonances could be readily identified and were in agreement with published spectra from purified CPS. In NCTC11168, the spectra clearly demonstrate the CPS anomeric protons from individual sugars and modifications, allowing simple screening of potential NCTC11168 capsular mutants to determine what residues are affected. Since capsular polysaccharides are the major

serodeterminant of the heat-stable typing scheme, HR-MAS NMR also provides a quick method of determining whether strains belong to the same Penner serogroup.

HR-MAS NMR also allowed us to examine structural capsule variants from the diverse NCTC11168 population. Phase variability of campylobacter capsule structures was noted as early as 1991 by Mills *et al.* when several strains showed serotyping differences after *in vitro* laboratory passage. The authors then observed differences in antibody response with typing sera after multiple *in vivo* samplings of the same strain. There are an abundance of variable bacterial sugar modifications mentioned in the literature, some of which have been recently summarized. A novel modification for *C. jejuni* NCTC11168 variant 2 was observed with -OP=O(NH₂)OMe on the 3-position of Gal/NAc. This structure could not be resolved in the previous study due to its low abundance in the wild type population. The phosphoramidate has not been described previously in nature and, it shows structural similarity to synthetic organophosphate insecticides. Future studies will determine the relevance of these phase variable modifications, identify the genes necessary for their biosynthesis, and examine the commonality of the phosphoramidate addition.

It is generally accepted that a single microorganism can give rise to a diverse population with very different virulence properties. However, sensitive methods for the structural analysis of bacterial populations have been limiting. As disclosed herein, CE-MS/MS can be used to examine the structure and variability in *C. jejuni* LOS. HR-MAS NMR has been used to investigate CPS structure, confirm serotype, demonstrate population variability, study the effect of mutagenesis, and detect *N*-linked glycoprotein sugars. Campylobacter has a large repertoire of variable surface glycans in addition to a conserved *N*-linked glycan. These studies have implications in vaccine development, provide possibilities for the induction of GBS following campylobacter enteritis, describe methods that can be adapted for the analysis of glycans from other important bacterial pathogens, expand the new field of metabolomics, and can provide more insight into the importance of bacterial LOS, capsules, and protein glycosylation allowing scientists to expand the discipline of glycomics beyond the gene complement and glycan structure.

The sequences of capsule biosynthetic loci from six *C. jejuni* strains were compared using a PCR amplification procedure based on the presence of highly conserved genes in this region. CPS structure prediction based on the analysis of these

sequences showed a good correlation with NMR and sugar analysis and with published data. The presence of additional genes in the *cps* regions suggests a potential for the biosynthesis of CPSs with modified structures. There is extensive duplication of glycosyltransferase genes in these loci resulting in approximately
5 double the number of transferases predicted by the structure.

Striking similarity between the *cps* regions of both the HS:23 and HS:36 serostrains and that of strain 81-176, which is of mixed HS:23/HS:36 serotype (Fig.6), suggests a common origin. Some variation in the respective CPS structures may be attributed to the presence of phase variable genes. Phase-variable expression of
10 methyl, ethanolamine, aminoglycerol and phosphoramidate groups on the CPS of strain NCTC11168 has been observed.

Sequence data was used in further comparison of the *cps* regions with that of the *C. jejuni* NCTC11168 strain. Both highly conserved and variable genes were found. The biosynthetic genes that are proximal to the transport- and assembly-related
15 *kps* genes were usually more conserved. The most conserved genes were the five to six genes near the *kpsC* gene. Interestingly, the study of *Streptococcus pneumoniae* *cps* loci also revealed a non-random variation of CPS-related genes, with the highest difference for those closest to the central region. The current data suggest that recombination events leading to variant forms of CPS of *C. jejuni* usually occur in the
20 middle of this region with the exception of the heptose biosynthetic genes (Fig. 10). In addition, most biosynthetic *cps* genes of serostrains HS:41 and strain NCTC11168, except for three genes in the middle, were found to have very low levels of similarity. This finding suggests that the *cps* regions of these strains are the most distantly related.

25 In addition to the mechanism of variation attributed to horizontal gene transfer, extensive intragenomic variation in the *cps* regions has been observed. Some genes, e.g. *cj1421* and *cj1422* in NCTC11168 share long regions of identity, which may have resulted from gene duplication. In other strains only one copy of these genes is present. Other genes may have also arisen from deletions resulting in
30 formation of hybrid genes. Such deletions/duplications may also play an important role in structural variation of CPSs. One interesting feature related to the mechanism of genetic variation was the finding of mosaic structure of some genes and their respective products. The N-terminal region of the Cj1440 homologue from serostrain

HS:19 (H19.11) revealed similarity to many *C. jejuni* glycosyltransferases, with the first 169 aa residues almost identical to the N-terminal residues of Cj1440 protein of NCTC11168. However, the C-terminus of HS19.11 revealed no similarity to the corresponding region of Cj1440, and resembled instead that of the Cj1438 glycosyltransferase. The finding supports the possibility of intra-cistron recombinations between the genes performing a similar function (e.g. encoding glycosyltransferases), which may result in altered substrate specificity and may contribute to antigenic variation of the CPS. This is similar to the observation in the LOS biosynthesis locus of the HS:10 serostrain which contains a β -1,4-*N*-acetylgalactosaminyltransferase (CgtA) and a β -1,3-galactosyltransferase (CgtB) that have diverged mostly in their C-termini when compared with the corresponding glycosyltransferases in the HS:19 serostrain {Gilbert, 2002 #181}. Furthermore, a number of genes in the *cps* region have homopolymeric tracts that may also contribute to variation.

The question remains, however, why so many different capsular structures are required for *C. jejuni*? Structural variation of CPSs is not restricted to *C. jejuni*. Other bacteria also developed various ways of changing cell surface properties through variation of CPSs. One remarkable example is *S. pneumoniae* where more than 90 different capsular serotypes have been described, the large majority of which are encoded by different cassettes at the same genomic locus. Extensive variation in the *cps* regions of *C. jejuni* adds to its arsenal of antigenic variation mechanisms involving cell surface structures. In addition, the requirement for CPS and variability of its structure may be dictated by changing host or environmental conditions. For example, colanic acid (exopolysaccharide) contributes to acid and heat tolerance in *E. coli*. Similarly, resistance of *C. jejuni* to heat treatment during food preparation may be attributed to certain cell surface located structures, including CPS. Variation in *C. jejuni* CPS structure may be a consequence of selective pressure in various environmental and *in vivo* conditions.

The discovery of the high conservation of some genes in the biosynthetic *cps* region along with the variation of others, serves as a basis for a PCR based typing procedure, which can provide a number of advantages over a classical Penner typing scheme. The limited number of antisera available for serotyping (usually a panel of 66 antisera) used in the standard Penner typing protocol results in up to 20% of strains

being untypeable. PCR amplification of the *cps* loci can allow for differentiation of these strains based on their potential of CPS production. For example, PCR analysis allowed detection of CPS-related genes in the untypeable strain X, known to produce a CPS. An advantage of a PCR-based typing scheme based on the sequences derived
5 from the *cps* regions is that it is based on the presence of the genes, rather than on their expression, which may be affected by a number of factors, including growth conditions. In addition, slight variation in the method of antigen preparation and conditions of passive hemagglutination may affect the results of typing using the classical Penner typing protocol. For example, the results of passive hemagglutination
10 depend on the origin of erythrocytes. Therefore, a PCR-based approach based on genetic difference in the *cps* regions can produce a more reliable and comprehensive typing scheme.

Multi-strain comparison of *C. jejuni* CPS loci has revealed a high conservation in genes involved in heptose biosynthesis and those flanking the *kps* regions,
15 particularly near *kpsC*. The findings suggest that CPS clusters are exchanged between *C. jejuni* and other bacteria and may in part be responsible for the structural variation observed. Other putative mechanisms of structural variation revealed here include gene duplication, deletion, recombination and contingency gene variation. Furthermore, there still remain a remarkably large number of genes with as yet
20 unknown function that may be involved in the biosynthesis of CPSs with modified structures. Analysis of the polysaccharides using NMR has provided novel CPS structural information, including the demonstration that the recently identified phosphoramidate modification is common to many *C. jejuni* CPSs, and where known, the predicted CPS structure showed good correlation with published structural data.
25 This "belt-and-braces" sequencing approach has provided an opportunity to test hypotheses formed regarding the structure of the capsular polysaccharides synthesized by the respective loci. In addition, these comparative studies form the basis for further work in the elucidation of heptose biosynthesis in *C. jejuni*. Future studies will confirm these correlations by mutagenesis, enzyme assays and analysis of general
30 carbohydrate metabolism. This study demonstrates the extensive variability of the CPS structural determinant in *C. jejuni* and underpins the genetic basis for Penner serotyping. The commonality of CPS-related heptose biosynthetic pathways among

bacteria and the presence of a mobile genetic element responsible for heptose biosynthesis in various strains of *C. jejuni* has also been described herein.

C. jejuni strains from different disease presentations and geographical locations were surveyed for the phosphoramidate (Table Xa). Examination of the closely related *Campylobacter coli* demonstrated that this modification is absent from this species (Table Xb) and also absent in other species sampled with the exception of one *C. fetus* isolate from a human with bacterial septicemia. Multiple colonies from selected mutants were analysed to demonstrate their potential role in phosphoramidate biosynthesis in NCTC11168. Mutagenesis of *cj1416c*, *cj1417c*, *cj1418c* and *cj1421c* resulted in loss of phosphoramidate while inactivation of the second copy of *cj1421* (ie. *cj1422c*) had no effect in NCTC11168 (Table Xc). However, based on the comparative sequencing results described above, *cj1422c* homologues are found in some phosphoramidate producing strains without *cj1421c* and appear to be responsible for the addition of phosphoramidate in these strains (for example the variable phosphoramidate observed in the HS:23, HS:36 and HS:23/36 group). Interestingly, mutagenesis of *cj1435c* and *cj1437c* resulted in loss of capsule and yet the phosphoramidate signal was still detected. These results suggest that this modification can either be added to additional structures and/or that we are detecting biosynthetic intermediates in the phosphoramidate pathway. A summary of the phosphoramidate survey results is shown in Table XI demonstrating that the phosphoramidate is common to 70% of *C. jejuni* isolates and 0% of *C. coli* isolates examined. Many strains were tested for the presence of genes *cj1416c*, *cj1421c* and *cj1422c* by PCR. The results summarizing the PCR probing and sequencing (which also includes *cj1417c* and *cj1418c*) are summarized in Table XIII and the primers used for probing are listed in Table XIV.

Phosphoramidate analysis demonstrated that *C. jejuni*, RM1221 is unable to produce phosphoramidate and this is consistent with this strain lacking the genes involved in its biosynthesis (Fig. 11). Comparison of the *C. jejuni* NCTC11168 capsule locus with other sequenced epsilon proteobacteria demonstrates that these strains also lack these genes and are likely unable to synthesize the phosphoramidate (Fig. 11). Thus, this structure appears to be unique to *C. jejuni* with rare exceptions in other *Campylobacter* species.

It is interesting to note that the phosphoramidate is added to capsules of strains with different serotypes and thus to different structures. A more thorough analysis of the capsule structures of two sero strains known to produce phosphoramidate, HS:1 and HS:19 was performed. As mentioned above, the HS:1 CPS structure consists of galactose and glycerol-phosphate. However, an unusual sugar that has not previously been described (see Table II) was detected. It is to this unusual sugar that the phosphoramidate is being attached (Fig. 12). More convincing results were obtained with HS:19 (Fig. 13). The phosphoramidate is attached at the 4-position of GlcNAc in contrast to NCTC11168 where the phosphoramidate is attached at the 3-position of GalNAc. As mentioned above, the phosphoramidate is detected in select mutants that lack the CPS. This suggests that the phosphoramidate is added to alternate structures or that intermediate forms can be detected. Occasionally additional phosphoramidate signals are observed during phosphate scans of capsulated *C. jejuni* isolates (Fig. 14). Thus, in addition to the possibilities mentioned, phosphoramidates may be attached to a varying CPS backbone which could lead to additional signals.

In order to determine whether the expression of phosphoramidate on *C. jejuni* CPS has any biological relevance, tissue culture and serum sensitivity assays were performed. Preliminary adherence and invasion assays comparing wildtype to the phosphoramidate mutants demonstrated that loss of phosphoramidate caused decreased adherence to CaCo-2 cells while invasion appeared unaffected (Fig. 15). To ensure that the differences in the tissue culture assays were not due to differences in motility of the mutants, motility assays were done comparing the wildtype strains with their respective mutants and demonstrated that all mutants had similar levels of motility compared to the parent (Fig. 16). Preliminary serum sensitivity assays comparing wildtype to the 1416-1 mutant indicated that expression of the phosphoramidate increases *C. jejuni* sensitivity to pooled human serum (Fig. 17). These results suggest that it would be beneficial for *C. jejuni* to suppress phosphoramidate expression while traveling to its optimal location and then expressing the phosphoramidate in order to efficiently bind to host cells.

¹ The abbreviations used are: Bac, bacillosamine, 2,4-diacetamido-2,4,6-trideoxy-D-glucopyranose; CE, capillary electrophoresis; CPMG, Carr-Purcell-Meiboom-Gill;

CPS, capsular polysaccharide; DIPSI-2, decoupling in the presence of scalar interactions; ESI-MS, electrospray ionization mass spectrometry; GBS, Guillain-Barré Syndrome; HR-MAS, high resolution magic angle spinning; LOS, lipooligosaccharides; LPS, lipopolysaccharide; MAS, magic angle spinning; NOESY, nuclear Overhauser effect spectroscopy; PVDF, polyvinylidene difluoride; TOCSY, total correlation spectroscopy; WURST-2, wideband, uniform rate, and smooth truncation; HMQC, heteronuclear multiple quantum correlation; HMBC, heteronuclear multiple bond correlation; NMR, nuclear magnetic resonance.

10 The presence of the identified phosphoramidate across most serotypes of *C. jejuni* makes it a useful target for attack on these cells, as well as for the identification of these cells.

 Molecules having a good binding affinity for this phosphoramidate (“binders”) can be employed, either alone or as conjugates or on the surface of liposomes or other
15 suitable cargo carriers or matrices to bind to *C. jejuni* cells. Where such molecules are functionally associated with a toxin or similar substance, they may be used to reduce *C. jejuni* viability or proliferation on a surface or in a solution, fluid, or semifluid of concern. Similarly, when such molecules are associated with a marker, such as an enzyme able to catalyze a colourimetric reaction, or a fluorescent or
20 radioactive marker, they can be used to identify and/or localize *C. jejuni*. One example of suitable binders for this purpose are antibodies having specificity for the phosphoramidate. Such antibodies may be single domain antibodies. In some cases only fragments of antibodies having the desired specificity will be employed. Such fragments may be expressed as part of a fusion protein with a “cargo” polypeptide of
25 interest.

 Another example of suitable binders are bacteriophages or portions thereof having a good affinity for the phosphoramidate. In some instances the phage particles or portions will also be capable of lysing the *C. jejuni* cells. Parts of interest can include any functional part. In some instances phagetail sheaths and/or tail spike proteins may
30 be employed.

 Other binder molecules can be identified by screening of materials for specific binding to the phosphoramidate. Thus, in one embodiment the invention provides a use

of the phosphoramidate in identifying compounds, or materials useful in identifying or reducing the viability of *C. jejuni*.

In some instances binders which recognize the phosphoramidate regardless of its sugar of attachment will be desired.

5 In some instances binders which recognize the phosphoramidate in association with one or more particular sugars of attachment will be desired.

Binders which specifically recognize the phosphoramidate structure unique to *Campylobacter* and no other phosphate compounds produced in nature are considered to be useful binders.

10 In an embodiment of the invention there is provided a method of modulating the adhesion of *C. jejuni* cells to a surface, comprising modulating the concentration of binders in the surrounding fluid. It will be understood that the surface may include a non-living material, cells, and/or cell-derived materials.

15

Table I

Chemical shifts^a (ppm) for the *C. jejuni* NCTC11168 variant 2 CPS

Unit		1	2	3	4	5	6	7	8
A	δ_C	106.2	81.2	70.7	84.0	63.0			
	δ_H	5.35	4.21	4.30	4.13	3.89, 3.70			
B	δ_C	99.0	73.1	73.8	76.4	72.7	171.3 ^c	53.9	61.3 ^b
	δ_H	5.14	3.94	4.10	3.93	4.38	8.32 ^c	4.03	3.72, 3.67 ^b
C	δ_C	106.4	62.5	79.6	82.2	78.3	61.9	174.9 ^c	22.9
	δ_H	5.13	4.27	4.88	4.48	3.97	3.82, 3.77	8.31 ^c	2.05
D	δ_C	98.1	72.2	73.8	70.2	72.3	79.5	63.1	60.4
	δ_H	5.61	3.53	3.72	3.56	4.08	3.80	3.86	3.55
E	δ_C	54.8							
	δ_H	3.75							

^a Measured at 600 MHz (¹H) in D₂O at 35⁰C (± 0.2 ppm error for d_C and ± 0.02 ppm for d_H).Internal acetone CH₃ resonance set at d_H 2.225 ppm and d_C 31.07 ppm.^b (CH₂)₂ of glycerol (8B and 9B).^c C=ONH. d_C from HMBC at 25⁰C. d_H from ¹H spectrum in 90% H₂O 10%D₂O at 25⁰C.

Table. II. Biosynthetic *cps* regions of various strains of *C. jejuni*.

Strain	Serotype	Size	Number of genes	Contingency genes	Sugar-phosphate nucleotidyl- transferases	Putative glycosyl- transferases	Number of different sug residues foun
NCTC11168	HS:2	34180	28	6	2	8	4
NCTC12517	HS:19	16727	13	2	1	4	2 [#]
G1	HS:1	15180	11	2	2	2	2 [#]
81-176	HS:23/36	24625	21	5	2	7	6
CCUG 10954	HS:23	24627	21	6	2	7	6
ATCC 43456	HS:36	24625	21	6	2	7	6
176.83	HS:41	34118	30	5	2	8	4

*According to structural analysis

[#]Additional uncharacterised labile substituent is present

Table III. Strain NCTC11168. In this and the following tables: nrdb – non-redundant database; contingency genes are shown in bold; the last three letters in the names of homologues, which indicate a gene product with the highest similarity score, are preceded by two letter codes for bacteria: Aa - *Aquifex aeolicus*, Af - *Archaeoglobus fulgidus*, At - *Aneurinibacillus thermoaerophilus*, Au - *Agrobacterium tumefaciens*, Bb - *Borrelia burgdorferi*, Bf - *Bacteroides fragilis*, Bh - *Bacillus halodurans*, Bj - *Bradyrhizobium japonicum*, Bs - *Bacillus subtilis*, Ca - *Clostridium acetobutylicum*, Cd - *Corynebacterium diphtheriae*, Ec - *E. coli*, Hi - *Haemophilus influenzae*, Kp - *Klebsiella pneumoniae*, Li - *Leptospira interrogans*, Ll - *Lactococcus lactis*, Mj - *Methanococcus jannaschii*, Mp - *Mycoplasma pneumoniae*, Mh - *Methanobacterium thermoautotrophicum*, Mt - *Mycobacterium tuberculosis*, No - *Nostoc* sp., Pa - *Pyrrococcus abyssi*, Pm - *Pasteurella multocida*, Sm - *Salmonella muenchen*, - Sp - *Streptococcus pneumoniae*, Sy - *Synechocystis* sp., Rm - *Rhizobium meliloti*, Tn - *Thermotoga neapolitana*, Ye - *Yersinia enterocolitica*, Yp - *Yersinia pseudotuberculosis*.¹ Function prediction of glycosyltransferases is based on either annotation by the Sanger Institute (for Cj1421, Cj1422, Cj1432, Cj1434, Cj1438 and Cj1440) or similarity to other glycosyltransferases of *C. jejuni* (for Cj1431 and Cj1442). These glycosyltransferases are labelled as conserved hypothetical without indicating E values.² According to {St Michael, 2002 #212}.

No	Gene	Homologues in nrdb	Putative function(s) of database homologue	Main substrate(s)	Main product(s)	Possible function, related to CPS biosynthesis
1	<i>cj1415</i>	PaCysC (4e-21)	adenylylsulfate kinase	adenylylsulfate	phosphoadenylylsulfate	unknown
2	<i>cj1416</i>	AtGalU (1e-26)	glucose-1-phosphate uridylyltransferase	UTP+glucose-1-phosphate	UDP-glucose	UDP-glucose pyrophosphorylase
3	<i>cj1417</i>	LiGuaA (3e-13)	GMP synthase	Xanthosine-5-phosphate	Guanosine-5-phosphate	unknown
4	<i>cj1418</i>	PpPpsA (2e-46)	pyruvate kinase	pyruvate	phosphoenolpyruvate	gluconeogenesis
5	<i>cj1419</i>	Mt-(4e-10)	methyltransferase	-	-	methyltransferase
6	<i>cj1420</i>	Mt-(9e-12)	methyltransferase	-	-	methyltransferase
7	<i>cj1421</i> ¹	conserved hyp	sugar transferase	-	-	sugar transferase
8	<i>cj1422</i> ¹	conserved hyp	sugar transferases	-	-	sugar transferase
9	<i>cj1423</i>	AtHddC (5e-55)	D-glycero-D-manno-heptose 1-phosphate	GTP+D-glycero-D-manno-heptose	GDP-D-glycero-D-manno-heptose	HddC, heptose 1-phosphate guanosyltransferase

		guanosyltransferase	1-phosphate		
10	<i>cj1424</i>	sedoheptulose 7-phosphate isomerase	sedoheptulose 7-phosphate	D-a-D-heptose-7-phosphate	GmhA, sedoheptulose 7-phosphate isomerase
11	<i>cj1425</i>	D-glycero-D-manno-heptose 7-phosphate kinase	D-a-D-heptose-7-phosphate	D-a-D-heptose-1,7-diphosphate	HddaA, heptose 7-phosphate kinase
12	<i>cj1426</i>	-	-	-	unknown
13	<i>cj1427</i>	UDP-glucose 4-epimerase	UDP-glucose	UDP-galactose	nucleotidyl-sugar epimerase
14	<i>cj1428</i>	GDP-fucose synthetase	GDP-4- keto-6-deoxy-D-mannose	GDP-fucose	GDP-heptose biosynthesis ²
15	<i>cj1429</i>	-	-	-	unknown
16	<i>cj1430</i>	dTDP-4-dehydro rhamnose 3,5-epimerase	dTDP-6-deoxy-D-xyl-4-hexulose	dTDP-6-deoxy-L-lyxo-4-hexulose	nucleotidyl-sugar epimerase
17	<i>cj1431</i> ¹	sugar transferase	-	-	sugar transferase
18	<i>cj1432</i> ¹	sugar transferase	-	-	sugar transferase
19	<i>cj1433</i>	-	-	-	unknown
20	<i>cj1434</i> ¹	sugar transferase	-	-	sugar transferase
21	<i>cj1435</i>	phosphoserine phosphatase	phosphoserine	serine	unknown
22	<i>cj1436</i>	histidinol phosphate transaminase	imidazole acetol phosphate	histidinol phosphate	unknown
23	<i>cj1437</i>	histidinol phosphate transaminase	imidazole acetol phosphate	histidinol phosphate	unknown
24	<i>cj1438</i> ¹	sugar transferase	-	-	sugar transferase
25	<i>cj1439</i>	UDP-galactopyranose mutase	UDP-D-galactopyranose	UDP-D-galactofuranose	UDP-GalNAc pyranose mutase ²
26	<i>cj1440</i> ¹	sugar transferase	-	-	sugar transferase
27	<i>cj1441</i>	UDP-glucose 6-dehydrogenase	UDP-glucose	UDP-glucuronate	Udg, UDP-glucose 6-dehydrogenase
28	<i>cj1442</i>	sugar transferase	-	-	sugar transferase

Table IV. Strain NCTC12517 (HS:19).

No	Homologues in NCTC11168	Homologues in nrdb	Putative function of database homologue	Main substrate(s)	Main product(s)	Possible function, related to biosynthesis
1	<i>cj1415</i> (2e-88)	AtCysC (5e-21)	adenylylsulfate kinase	adenylylsulfate	phosphoadenylylsulfate	unknown
2	<i>cj1416</i> (9.3e-135)	AtGalU (4e-17)	glucose-1-phosphate uridylyltransferase	UTP+glucose-1-phosphate	UDP-glucose	sugar-1-phosphate nucleotidyltransferase
3	<i>cj1417</i> (4.7e-108)	CaGuaA (3e-13)	GMP synthase	Xanthosine-5-phosphate	Guanosine-5-phosphate	unknown
4	<i>cj1418</i> (0.0)	CaPpsA (2e-46)	pyruvate kinase	pyruvate	phosphoenolpyruvate	gluconeogenesis
5	<i>cj1419</i> (3.0e-138)	Mt-(4e-10)	methyltransferase	-	-	methyltransferase
6	<i>cj1420</i> (2.0e-141)	Mt-(9e-12)	methyltransferase	-	-	methyltransferase
7	<i>cj1421</i> ¹ (2.1e-153)	conserved hyp	sugar transferase	-	-	sugar transferase
8	<i>cj1440</i> (4e-118)	conserved hyp	sugar transferase	-	-	sugar transferase
9	<i>cj1435</i> (5e-77)	CaSerB (8e-11)	phosphoserine phosphatase	phosphoserine	serine	unknown
10	<i>cj1437</i> (7e-111)	MJHisC (5e-37)	histidinol phosphate transaminase	imidazole acetol phosphate	histidinol phosphate	unknown
11	<i>cj1438</i> (9e-146)	conserved hyp	sugar transferase	-	-	sugar transferase
	<i>cj1440</i> (2e-111)	-	-	-	-	sugar transferase
12	<i>cj1441</i> (4e-120)	PmUdg (1e-124)	UDP-glucose 6-dehydrogenase	UDP-glucose	UDP-glucuronate	Udg, UDP-glucose 6-dehydrogenase
13	<i>cj1442</i> (1e-113)	conserved hyp	sugar transferase	-	-	sugar transferase

Table V. Strain G1 (HS:1).

No	Homologues in	Homologues in	Putative function of database homologue	Main substrate(s)	Main product(s)	Possible function, related to CPS biosynthesis
	NCTC11168	nrdB				
1	<i>ej1415</i> (9.1e-89)	AfCysC (5e-21)	adenylylsulfate kinase	adenylylsulfate	phosphoadenylylsulfate	unknown
2	<i>ej1416</i> (3.1e-134)	AfGalU (1e-17)	glucose-1-phosphate uridylyltransferase	UTP+glucose-1-phosphate	UDP-glucose	sugar-1-phosphate
3	<i>ej1417</i> (2.5e-109)	BbGuaA (2e-07)	GMP synthase	Xanthosine-5-phosphate	Guanosine-5-phosphate	nucleotidyltransferase
4	<i>ej1418</i> (0.0)	PpPpsA (2e-46)	pyruvate kinase	pyruvate	phosphoenolpyruvate	unknown
5	<i>ej1419</i> (2e-139)	Mt-(4e-10)	methyltransferase	-	-	gluconeogenesis
6	<i>ej1420</i> (7e-144)	Mt-(9e-12)	methyltransferase	-	-	methyltransferase
7	<i>ej1421</i> ¹ (1e-151)	conserved hyp	sugar transferase	-	-	methyltransferase
8	<i>ej1432</i> (2e-04)	conserved hyp	sugar transferase	-	-	sugar transferase
9	<i>ej1127</i> (6e-20, C-term only)	BsTagF (4e-38)	sugar transferase (CDP glycerophosphotransferase)	-	-	sugar transferase
10	-	hyp	-	-	-	CDP glycerophosphotransferase
11	<i>ej1150</i> (1e-14)	EcTagD (8e-41)	glycerol-3-phosphate cytidylyltransferase	CTP + glycerol 3-phosphate	CDP-glycerol	unknown glycerol-3-phosphate cytidylyltransferase

Table VI. Strains 81-176 (HS:23/HS:36), CCUG 10954 (HS:23) and ATCC 43456 (HS:36).

No	Homologues	in	Homologues in nrdb ¹	% Identity (aa)				Putative function of database homologue	Main substrate(s)	Main product(s)	Possible function, related to CPS
				81-176 vs CCUG 10954	81-176 vs ATCC 43456	CCUG 10954 vs ATCC 43456					
1	<i>cj1415</i> (2e-88)		CaCysC (3e-21)	97.1	95.9	98.8	adenylylsulfate kinase	adenylylsulfate	phosphoadenylylsulfate	unknown	biosynthesis
2	<i>cj1416</i> (7e-135)		PpGalU (3e-26)	98.4	97.2	98.8	glucose-1-phosphate uridylyltransferase	UTP+glucose-1-phosphate	UDP-glucose	sugar-1-phosphate nucleotidyltransferase	
3	<i>cj1417</i> (5e-109)		CaGuaA (3e-13)	95.0	95.5	98.5	GMP synthase	Xanthosine-5-phosphate	Guanosine-5-phosphate	unknown	
4	<i>cj1418</i> (0.0)		AaPpsA (2e-07)	97.7	96.9	98.5	pyruvate kinase	pyruvate	phosphoenolpyruvate	gluconeogenesis	
5	<i>cj1419</i> (2e-136)		Mt-(4e-10)	97.6	97.6	98.4	methyltransferase	-	-	-	methyltransferase
6	<i>cj1420</i> (2e-143)		Mt-(9e-12)	98.5	100	98.5	methyltransferase	-	-	-	methyltransferase
7	<i>cj1422</i> ¹ (1e-120)		conserved hyp	94.8	98.8	93.9	sugar transferase	-	-	-	sugar transferase
8	<i>cj1423</i> (6e-106)		AtHddC (1e-54)	86.6	94.6	87.9	D-glycero-D-manno-heptose 1-phosphate guanosyltransferase	GTP+D-glycero-D-manno-heptose 1-phosphate	GDP-D-glycero-D-manno-heptose	HddC, heptose 1-phosphate guanosyltransferase	
9	<i>cj1424</i> (3e-103)		AtGmhA (7e-78)	98.0	97.0	96.0	sedoheptulose 7-phosphate isomerase	sedoheptulose 7-phosphate	D-a-D-heptose-7-phosphate	GmhA, sedoheptulose 7-phosphate isomerase	
10	<i>cj1425</i> (1e-177)		AtHddA (1e-105)	98.2	100	98.2	D-glycero-D-manno-heptose 7-phosphate kinase	D-a-D-heptose-7-phosphate	D-a-D-heptose-1,7-diphosphate	HddA, heptose 7-phosphate kinase	
11	<i>cj1427</i> (1e-164)		LiGalE (1e-33)	98.7	99.7	98.7	UDP-glucose 4-epimerase	UDP-glucose	UDP-galactose	nucleotidyl-sugar epimerase	

12 ²	<i>cj1319</i> (1e-26)	YpDmhA (1e-151)	97.7	99.7	98.0	GDP-mannoheptose-4,6 dehydratase	GDP-D-mannoheptose	GDP-4-dehydro-6-deoxy-D- mannoheptose	DmhA, GDP-mannoheptose-4,6 dehydratase
13	<i>cj1428</i> (2e-108)	Bffcl (7e-109)	99.7	99.4	99.7	GDP-fucose synthetase	GDP-4- keto-6-deoxy-D- mannose	GDP-fucose	GDP-4-keto-6-deoxy-D-sugar-3,5,5- epimerase-4-reductase (putative heptose biosynthesis - by similarity to Cj1428 ²)
14	<i>cj1429</i> (6e-147)	hyp	100	100	100	-	-	-	unknown
15	<i>cj1430</i> (2e-83)	SeRfbC (2e-21)	100	100	100	dTDP-4-dehydro rhamnose 3,5-epimerase	dTDP-6-deoxy-D-xyl-4- hexulose	dTDP-6-deoxy-L-lyxo-4-hexulose	nucleotidyl-sugar epimerase
16	<i>cj1431</i> (7.6e-49)	conserved hyp	100	100	100	sugar transferase	-	-	sugar transferase
17	<i>cj1138</i> (6.2e-106)	conserved hyp	96.0	98.6	94.9	sugar transferase	-	-	sugar transferase
18	<i>cj1440</i> (3e-107)	conserved hyp	97.4	100	97.4	sugar transferase	-	-	sugar transferase
19	<i>cj1434</i> (2e-07)	conserved hyp	99.8	100	99.8	sugar transferase	-	-	sugar transferase
20	<i>cj1422</i> (1e-140)	conserved hyp	94.7	98.8	93.5	sugar transferase	-	-	sugar transferase
21	<i>cj1442</i> (2e-113)	conserved hyp	99.0	99.6	98.9	sugar transferase	-	-	sugar transferase

¹Similarity is calculated using the sequence from strain 81-176.

²ORF #12 is phase-variable in sero strains HS:23 and HS:36 but not in strain 81-176.

Table VII. Strain 176.83 (HS:41).

No	Homologues in NCTC11168	Homologues in nrdb	Putative function of database homologue	Main substrate(s)	Main product(s)	Possible function, related to CFS biosynthesis
1	<i>cj1231</i> (4e-14)	EcKefB (4e-29)	Na ⁺ /H ⁺ antiporter	-	-	unknown
2	<i>cj0392</i> (3e-27)	CaPyk (2e-36)	pyruvate kinase	ATP + pyruvate	ADP + phosphoenolpyruvate (final step in glycolysis)	unknown
3	<i>cj1609</i> (1e-27)	BsSat (6e-74)	sulfate adenylyltransferase	ATP + sulfate	diphosphate + adenylylsulfate (Cys biosynthesis)	unknown
4	<i>cj1415</i> (2e-56)	AtCysC (7e-22)	adenylylsulfate kinase	adenylylsulfate	phosphoadenylylsulfate	unknown
5	<i>cj1422</i> (7e-96)	conserved hyp	sugar transferase	-	-	sugar transferase
6	<i>cj1431</i> (9e-23)	conserved hyp	sugar transferase	-	-	sugar transferase
7	<i>cj1423</i> (4e-103)	AtHddC (4e-53)	D-glycero-D-manno-heptose 1-phosphate guanosyltransferase	GTP+D-glycero-D-manno-heptose 1-phosphate	GDP-D-glycero-D-manno-heptose	HddC, heptose 1-phosphate guanosyltransferase
8	<i>cj1424</i> (3e-101)	AtGmhA (7e-76)	sedoheptulose 7-phosphate isomerase	sedoheptulose 7-phosphate	D-a-D-heptose-7-phosphate	GmhA, sedoheptulose 7-phosphate isomerase
9	<i>cj1300</i> (4e-06)	conserved hyp	sugar transferase	-	-	sugar transferase
10	<i>cj1425</i> (5e-174)	AtHddA (1e-106)	D-glycero-D-manno-heptose 7-phosphate kinase	D-a-D-heptose-7-phosphate	D-a-D-heptose-1,7-diphosphate	HddA, heptose 7-phosphate kinase
11	<i>cj1319</i> (2e-27)	YpDmhA (e-155)	GDP-mannoheptose-4,6 dehydratase	GDP-D-mannoheptose	GDP-4-dehydro-6-deoxy-D-mannoheptose	DmhA, GDP-mannoheptose-4,6 dehydratase
12	<i>cj1428</i> (6e-106)	BfFcl (1e-116)	GDP-4-keto-6-deoxy-D-mannose-3, 5-epimerase-4-reductase (GDP-fucose synthetase)	GDP-4-dehydro-6-deoxy-D-mannose	GDP-fucose	GDP-4-keto-6-deoxy-D-sugar-3,5-epimerase-4-reductase (putative fucose or heptose biosynthesis - by similarity to Cj1428 ²)
13	<i>cj1428</i> (2e-104)	EcFcl (1e-111)	GDP-4-keto-6-deoxy-D-mannose-3, 5-epimerase-4-reductase (GDP-fucose)	GDP-4-dehydro-6-deoxy-D-mannose	GDP-fucose	GDP-4-keto-6-deoxy-D-sugar-3,5-epimerase-4-reductase (putative fucose or heptose)

			synthetase)				biosynthesis - by similarity to Cj1428 ²)
14	<i>cj1430</i> (9e-82)	SmRfC (4e-22)	dTDP-4-dehydroglucose 3,5-epimerase	dTDP-4-dehydro-6-deoxy-D-glucose	dTDP-4-dehydro-6-deoxy-L-mannose.	nucleotidyl-sugar epimerase	
15	<i>cj1439</i> (6e-03)	CaGlf (1e-09)	UDP-galactopyranose mutase	UDP-D-galactopyranose	UDP-D-galacto-1,4-furanose	nucleotidyl-sugar pyranose mutase	
16	<i>cj1431</i> (1e-60)	conserved hyp	sugar transferase	-	-	sugar transferase	
17	<i>cj1439</i> (3e-52)	CaGlf (1e-66)	UDP-galactopyranose mutase	UDP-D-galactopyranose	UDP-D-galacto-1,4-furanose	nucleotidyl-sugar pyranose mutase	
18	<i>cj1319</i> (4e-14)	SmRfBj (2e-21)	CDP-abequose synthase	CDP-4-keto 3,6-dideoxyglucose	CDP-abequose	nucleotidyl-sugar dehydratase	
19	<i>cj1430</i> (1e-14)	SyRfC (7e-39)	dTDP-4-dehydroglucose 3,5-epimerase	dTDP-4-dehydro-6-deoxy-D-glucose	dTDP-4-dehydro-6-deoxy-L-mannose.	nucleotidyl-sugar epimerase	
20	<i>cj1319</i> (8e-18)	LiRfG (3e-86)	CDP-glucose 4,6-dehydratase	CDP-glucose	CDP-4-dehydro-6-deoxy-D-glucose	nucleotidyl-sugar dehydratase	
21	<i>cj1423</i> (2e-13)	YeDdhA (8e-87)	glucose-1-phosphate cytidyltransferase	CTP + D-glucose 1-phosphate	CDP-glucose	sugar-1-phosphate nucleotidyltransferase	
22	<i>cj1438</i> (6e-15)	conserved hyp	sugar transferase	-	-	sugar transferase	
23	<i>cj1438</i> (7e-20)	conserved hyp	sugar transferase	-	-	sugar transferase	
24	<i>cj1439</i> (3e-65)	MhGlf (1e-106)	UDP-galactopyranose mutase	UDP-D-galactopyranose	UDP-D-galacto-1,4-furanose	nucleotidyl-sugar pyranose mutase	
25	<i>cj1131</i> (1e-73)	NoGalE (1e-95)	UDP-glucose 4-epimerase	UDP-glucose	UDP-galactose	sugar epimerase	
26	<i>cj1441</i> (2e-41)	AuUdg (1e-115)	UDP-glucose 6-dehydrogenase	UDP-glucose	UDP-glucuronate	Udg, UDP-glucose 6-dehydrogenase	
27	<i>cj1319</i> (7e-34)	BjGalE (1e-110)	UDP-glucose 4-epimerase	UDP-glucose	UDP-galactose	nucleotidyl-sugar pyranose mutase	
28	<i>cj1138</i> (1e-107)	conserved hyp	sugar transferase	-	-	sugar transferase	
29	<i>cj1434</i> (3e-79)	conserved hyp	sugar transferase	-	-	sugar transferase	
30	<i>cj1442</i> (2e-282)	conserved hyp	sugar transferase	-	-	sugar transferase	

Table VIII. Analysis of contingency genes in the biosynthetic *cps* region

Strains	Genes ¹						
	1420	1421	1422	1426	1429	1437	
G1 (HS:1)	G10/G9	G9	np ³	np	np	np	
NCTC12500 (HS:1)	G9/G10	G9/G10	np	np	np	np	
NCTC11168 (HS:2) ²	G9/G10	G8/G9/G10	G9/G10	G10/G11	G9/G10	G9	
11168H (HS:2) ³	G9	G9	G9	G10	G10	G9	
NCTC12501 (HS:2)	G9/G10	G9/G10	G9	G10	G10/G11/G12	G9	
NCTC12501 (HS:3)	np ³	np	np	np	np	np	
NCTC12561 (HS:4)	G9	G8	np	np	np	np	
81116 (HS:6)	np	np	np	np	np	np	
NCTC12509 (HS:10)	np	np	np	np	np	np	
NCTC12517 (HS:19)	G9	G9	np	np	np	G9	
X (untypeable)	G9	G9	np	np	np	np	
Gene "ON"	G9	G9	G9	G10	G10	G9	
"ON"/"OFF"	63%	67%	90%	91%	90%	100%	
in NCTC11168 ⁴							

¹Genes detectable with ak149/ak150 primers (see Experimental Procedures)²{Parkhill, 2000 #24}³Single colony hypermotile isolate of NCTC11168⁴Ratio of "ON"/"OFF" genes in the sequenced genome of NCTC11168 {Parkhill, 2000 #24}. ⁵No detectable PCR product

Table IX. Comparison of the contingency genes in the biosynthetic *cps* region of serostrains HS:23, HS:36 and strain 81-176:

Genes						
Strains	ORF No.6 (<i>ej1420</i>) ¹	ORF No.7 (<i>ej1422</i>)	ORF No.12 (<i>ej1319</i>)	ORF No.14 (<i>ej1429</i>)	ORF No.17 (<i>ej1138</i>)	ORF No.20 (<i>ej1422</i>)
81-176 (HS:23/HS:36)	G9 ²	G9	(on) ³	G9/G10	G9	G8/G9/G10
CCUG 10954 (HS:23)	G8/G9	G8	G9	G8/G9	G9	G8
ATCC 43456 (HS:36)	G8/G9	G8/G9	G9/G10	G9/G10	G9	G9/G10
Gene "ON"	G9	G9	G9	G9	G9	G9

¹ Homologue in strain NCTC11168.² The number in bold indicates the most frequent variant.³ ORF No.12 is not phase-variable in strain 81-176.

Table X The commonality of the phosphoramidate moiety on *C. jejuni* capsular polysaccharides as determined by the P-filter method using HR-MAS NMR. Structural work has indicated that the phosphoramidate is found on position-3 of GalNAc in NCTC11168, on position-4 of GlcNAc of HS:19 and on a unique deoxy sugar in HS:1. Empty boxes indicate experiments that have not been performed. The Penner serotype refers to the heat-stable (HS) type. Phage types (PT) indicated as RDNC refer to 'react with phages but do not conform to designated type'. Table a) includes *C. jejuni* wildtype strains, b) includes *Campylobacter* species other than *C. jejuni* and c) includes results for mutants in genes affecting phosphoramidate biosynthesis or accumulation. Note that V4 represents variant 2 described above while V9 represents variant 3 above.

a)

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Penner serotype	Phage Type
NCTC11168 (UK)	variant of NCTC11168	2 moieties clustered	HS2	
NCTC11168 version 1 (V1D0)	variant of NCTC11168	2 moieties	HS2	
NCTC11168 version 26 (NRC)	variant of NCTC11168	2 minor moieties clustered (x3?)	HS2	
NCTC11168 version 26 -V1	individual colony of Version 26	2 moieties clustered		
NCTC11168 version 26 -V4	individual colony of Version 26	predominantly 1 moiety		
NCTC11168 version 26 -V6	individual colony of Version 26	none		
NCTC11168 version 26 -V9	individual colony of Version 26	2 moieties clustered		
HS:19	enteritis	3 variable moieties clustered	HS19	
HS:1	enteritis	1 moiety	HS1	

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Penner serotype	Phage Type
81-176	enteritis	1 moiety	HS23/36	
HS:23	enteritis	none	HS23	
HS:36	enteritis	1 moiety	HS36	
Cura 235	enteritis	1 or 2 moieties clustered		
HS:41	enteritis	none	HS41	
RM 1244	human isolate	2 moieties (1 large, 1 small)		
Cj 60	watery diarrhea Brazil	2 or 3 moieties clustered	HSUT	
Cj 34007	septicemia	no, other phosphodiester x2? Smaller couplings at 5.5	HS18	2
Cj 47886	septicemia	2 moieties clustered	HSUT	36
Cj 52472	septicemia	2 moieties shifted apart	HSUT	1
Cj 47439	septicemia	none	HS67	1
Cj 53250	septicemia	none	HS60	2
Cj 44119	septicemia	none	HS18	2
Cj 43983	septicemia	2 moieties apart (decoupled spectrum shows possible third)	HS50	6
Cj 46979	septicemia	3 moieties clustered and apart	HS50	6
Cj 33084	asymptomatic	none	HS35	1
Cj 31485	asymptomatic	1 moiety	HS1	2
Cj 32799	asymptomatic	1 moiety	HS64G	5
Cj 32787	asymptomatic	2 moieties apart	HS18	RDNC
Cj 33106	asymptomatic	2 moieties clustered	HS23G	1
Cj 31481	asymptomatic	1 or 2 moieties clustered	HS17G	76

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Penner serotype	Phage Type
Cj 31467	asymptomatic	none	HS18	RDNC
MF 20	Miller Fisher	1 or 2 moieties clustered		
MF 7	Miller Fisher	1 or 2 moieties clustered		
MF 6	Miller Fisher	2 moieties clustered		
MF 8	Miller Fisher	1 moiety		
GB 2	Guillain-Barré syndrome	1 or 2 moieties clustered, small		
Cj G1	Guillain-Barré syndrome	1 moiety	HS1	
GB 11	Guillain-Barré syndrome	2 moieties clustered	HS2	
OH 4384	Guillain-Barré syndrome	none	HS19	
Cura 276	Guillain-Barré syndrome	2 moieties clustered (decoupled spectrum shows possible third)		
Cj 178	retail meat isolate	none		
Cj 198	retail meat isolate	2 or 3 moieties clustered		
Cj 160	retail meat isolate	none		
Cj 120	retail meat isolate	none		
Cj 150	retail meat isolate	none		
Cj 212	poultry isolate	none		
Cj 206	poultry isolate	none		

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Penner serotype	Phage Type
Cj 207	poultry isolate	2 or 3 moieties clustered		
Cj 220	poultry isolate	none		
Cj 230	poultry isolate	none		
RM 1221	poultry isolate	none	HS53	
RM 1184	poultry isolate	1 moiety		
RM 1216	poultry isolate	none		
RM 1285	poultry isolate	1 moiety	HS19	
RM 1850	poultry isolate	1 moiety		
Cj 11974	poultry isolate	2 moieties clustered (1 peak much stronger than other)	HS44	2
Cj 12567	poultry isolate	3 or 4 moieties clustered	HS2	36
Cj 11919	poultry isolate	2 moieties clustered	HS2	35
RM 3668	canada goose dropping	1 moiety		
Cj 12241	ovine	predominantly 1 moiety, but possibly more	HS50	6
Cj 11848	animal isolate	1 moiety		
NCTC12658	phage propagating	1 large		
NCTC12659	phage propagating	1 large		
NCTC12660	phage propagating	1 large		
NCTC12661	pigeon isolate, phage propagating	1 large		
NCTC12663	phage propagating	none		

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Penner serotype	Phage Type
NCTC12664	phage propagating	1 large		
NCTC12665	phage propagating	3 moieties (2 major and 1 minor)		

b)

<i>Campylobacter</i> species	Disease caused	POMe-filter	Penner serotype	Phage Type
<i>C. coli</i> HS:30		none	HS30	
<i>C. coli</i> 167		none		
<i>C. coli</i> ROH1		none		
<i>C. coli</i> ROH2		none		
<i>C. coli</i> 423		none		
<i>C. coli</i> 345		none		
<i>C. coli</i> 456		none		
<i>C. coli</i> 11		none		
<i>C. coli</i> 253		none		
<i>C. coli</i> 145		none		
<i>C. coli</i> 525		none		
<i>C. coli</i> 255		none		
<i>C. coli</i> 463		none		
<i>C. coli</i> 700		none		
<i>C. coli</i> RM2228		none	HS34	
<i>C. coli</i> RM1515 (ATCC33559)	pig feces	none		
<i>C. coli</i> RM1891	poultry isolate	none		
<i>C. coli</i> RM1900	swine farm	none		
<i>C. hominus</i> CH3NC1		none		
<i>C. fetus</i> 50618	septicemia	1 moiety		
<i>C. fetus</i> fetus		no, other weak phosphodiester at 4.17ppm with smaller couplings		
<i>C. fetus</i> venerealis ATCC19438		none		

c)

<i>C. jejuni</i> mutants	Parent/Notes	POMe-filter
Cj 1416-1	UK111168H	none
Cj 1416-2	UK111168H	none
Cj 1416-3	UK111168H	none
Cj 1417-1	UK111168H	none
Cj 1417-2	UK111168H	none
Cj 1417-3	UK111168H	none
Cj 1418-1	UK111168H	none
Cj 1418-2	UK111168H	none
Cj 1418-3	UK111168H	none
Cj 1421-3	V26 - variant 4 / knock-out of Cj1421c while Cj1422c naturally off	none
Cj 1422-2	V26 - variant 4 / knock-out of Cj1422c while Cj1421c naturally off	none
Cj 1422-3	same as 1422-2	none
Cj 1422-4	same as 1422-2	none
Cj 1435-1	UK111168H	present, very weak
Cj 1435-2	UK111168H	2 moieties, weak
Cj 1435-3	UK111168H	2 moieties, weak
Cj 1437-1	UK111168H	2 moieties clustered
Cj 1437-2	UK111168H	present, weak
Cj 1437-3	UK111168H	present, very weak

Table XI A summary of the distribution of the phosphoramidate on the capsular polysaccharide of *Campylobacter* strains isolated from a variety of animal sources and geographical locations.

<i>Campylobacter jejuni</i>			
5	Asymptomatic	(UK)	5/7
	Enteritis	(Various locations)	8/10
	Enteritis (phage propagating)	(NCTC)	5/6
	GBS	(Rotterdam / Japan)	4/5
	MF	(Rotterdam)	4/4
10	Septicemia	(UK)	5/8
	Poultry and retail meat	(Canada)	2/10
	Chicken and goose droppings	(US)	4/6
	Pigeon (phage propagating)	(NCTC)	1/1
	Animal	(UK)	5/5
15	Total		43/62 = 70%
Other Species			
20	<i>Campylobacter coli</i>	(Canada, US)	0/18 = 0%
	<i>Campylobacter fetus</i>	(UK)	1/1*
	<i>Campylobacter fetus fetus</i>	(NCTC)	0/1
	<i>Campylobacter fetus venerealis</i>	(ATCC)	0/1
	<i>Campylobacter hominus</i>	(UK)	0/1
	Total strains examined		84
25	*only non <i>C. jejuni</i> examined with phosphoramidate – this strain was isolated from a human with septicemia		

Table XIII: The commonality of genes involved in phosphoramidate biosynthesis. The presence of *Cj1416c*, *Cj1417c*, *Cj1418c*, *Cj1421c* and *Cj1422c* homologues was determined using sequencing and gene-specific polymerase chain reactions using primers found in Table XIV. The on/off status of *Cj1421c* and *Cj1422c*, if known, is indicated as these genes are phase-variable, and was determined by sequencing. Empty boxes indicate that the experiment was not done.

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Cj1416c	Cj1417c	Cj1418c	Cj1421c	Cj1422c
NCTC11168 (UK)	variant of NCTC11168	2 moieties clustered	present	present	present	present - on	present - on
NCTC11168 version 1 (VDO)	variant of NCTC11168	2 moieties	present	present	present	present - on	present - on
NCTC11168 version 26 (NRC)	variant of NCTC11168	2 minor moieties clustered (x3?)	present	present	present	present - majority off	present - majority on
NCTC11168 version 26 -V1	individual colony of Version 26	2 moieties clustered	present	present	present	present	present
NCTC11168 version 26 -V4	individual colony of Version 26	predominantly 1 moiety	present	present	present	present - on	present - on
NCTC11168 version 26 -V6	individual colony of Version 26	none	present	present	present	present - off	present - off
NCTC11168 version 26 -V9	individual colony of Version 26	2 moieties clustered	present	present	present	present	present
HS:19	enteritis	3 variable moieties clustered	present	present	present	present - on	diverged
HS:1	enteritis	1 moiety	present	present	present	present	diverged
81-176	enteritis	1 moiety	present	present	present	diverged	two copies - on and majority on
HS:23	enteritis	none	present	present	present	diverged	two copies - both off
HS:36	enteritis	1 moiety	present	present	present	diverged	two copies - on and off
Cura 235	enteritis	1 or 2 moieties clustered	present				
HS:41	enteritis	none	diverged	diverged	diverged	diverged	present - on

<i>C. jejuni</i> Strain	Disease caused	POMe-filter	Cj1416c	Cj1417c	Cj1418c	Cj1421c	Cj1422c
Cj 52472	septicemia	2 moieties shifted apart	present			present - off	present - on
Cj 47439	septicemia	none	diverged			present - on	present
Cj 46979	septicemia	3 moieties clustered and apart	present			present - on	diverged
Cj G1	Guillain-Barré syndrome	1 moiety	present	present	present	present - on	diverged
OH 4384	Guillain-Barré syndrome	none				diverged	diverged
Cj 160	retail meat isolate	none	diverged			diverged	diverged
Cj 206	poultry isolate	none	present			diverged	diverged
RM 1221	poultry isolate	none				present - on	diverged

Table XIV The sequences of primers used for the identification and characterization of *Cj1416c*, *Cj1421c* and *Cj1422c*. The chromosomal location represents the location of the primer on the forward strand in the genome of NCTC11168.

Gene	Primer Name	Sequence	Chromosomal Location
Cj1416c	Cj1416cR1	ATGAATGCAATTATCTTAGCAGCAG	1348851 - 1348876
Cj1416c	Cj1416cF758	TTTCAATTTGCACCTCTAAATCAC	1349585 - 1349610
Cj1421c	Cj1421cR7	AACCCAAATTCAGCCATAGAAAGAG	1356023 - 1356047
Cj1421c	Cj1421cF1839	CTAAATATCACCATCCAACCTCCTTG	1354215 - 1354239
Cj1421c	1421 spec F1	TTGCCTAATACTTTGGAAGG	1354907 - 1354926
Cj1421c	1421 spec B1	AAAAACCCACTGAACTATCC	1354521 - 1354540
Cj1421c	1421-R	AAAGGAAAAAACCCATGCTCA	1356066 - 1356047
Cj1421c	1421-F	TGGGTATTTAAGTTGGGGAAA	1355853 - 1355873
Cj1421c	1421LR	GACCCATCATGGTATGGCAATC	1355121 - 1355142
Cj1421/22	1421-22specB4	AACCTGTATTGATTTTGCTTAGAGATCC	1356997 - 1357024
Cj1421/22	1421/22specF2	TCCATAAAAACGATTGCGGC	1354880 - 1354899
Cj1422c	Cj1422cR7	AACCCAAATTCAGCCATAGAAAGAG	1357956 - 1357980
Cj1422c	Cj1422cF1869	AGAAAATGATAAAGATGGGTTTTGA	1356117 - 1356141
Cj1422c	1422 spec B13	GGATTTTTTTTGCATACAAGTGGAG	1357226 - 1357250
Cj1422c	1422 spec F1	GAGCTTTGATACTTTCAAGAGATGG	1356806 - 1356830
Cj1422c	1422-R	TTGGTGTGCCTGAGGATTATT	1358059 - 1358079
Cj1422c	1422-F	TGGGTATTTAAGTTGGGGAAA	1357785 - 1357805
Cj1422c	1422LR	GAGAGATTTTTTAAAAGAAGTTGGTG	1357191 - 1357215
Cj1422c	1422LF	CTGGGGAAGAAAAATTTCAGTATTTTG	1356639 - 1356665

Table XV

Part A

cj1421c:

5 MLNPNSAIERVKNHLAYKLGQTVIEHRHNGGGYIALFKKLYKIKKQHKKEQ
 KIYQQIIQV
 FPQLKYPSLETCSYNEALRCKFHLSYMIGEVLKAYQNWYKGGGFKLKNN
 IKKANKEFQ
 10 IFREILKEFKELNGETLKAIQDNKQLFLKEFPRIKNILKTHQDYQPILDNI
 FHNFNFIK
 NFDLIEEWLLSDDFKEKYKKENHPYPSLLDPKKLNDENEKINYHNIPAELE
 WKMNLPPLP
 NYEFMWFFSHGAGFTLGQFFYHLFKINILDYFCGGDGDIRYYKFYNKLE
 LKDKRNIIT
 15 INDIDPSWYGNQHKRDKLFSSFQKITPILFQIRDPIELIKHAYGRKWGNNL
 AKTKEFDLS
 YQFNDIITEVEVYNYNLENTLEGQRPQSFLWKSLECFDKFNDCFYLDISK
 IRGEETIHT
 LNYLSNKFNLKQIKINDKEFVTKSYFKGNLYFLLPLTLYLKEDLNTNIPN
 20 KKINKNNSL
 IININFFQNDNNLNLSELSILDMDSSVGFYIDKQDYNKLKNDISIFYKQV
 IDYLRNFAY
 ELKNRIQIEEDLMLKVEDVLRHLYMNKNARVSAKNILDEELVYIKQHRPDI
 VASWKYYQE
 25 FEKMCKELDGGDI

Part B

cj1422c:

30 MLNPNSAIERVKNHLAYKLGQTVIEHRHNGGGYIALFKKLYKIKKQHKKEQ
 KIYQQIIQV
 FPQLKYPSLETCSYNEALRCKFHLSYMIGEVLKAYQNWYKGGGFKLKNN
 IKKANKEFQ
 35 IFREILKEFKELNGETLKAIQDNKQLFLKEFPRIKNILKTHQDYQPILDNI
 FHNFNFIK
 NFDLIEEWLLSDDFKEKYKKENHPYPSLLDPKKLNDENEKINYHNIPAELE
 WKMNLPPLP
 NYEFVGGFFLHTSGEKAMERFLKEVGVLIGAFGYEDGKRYISIFNFLISEA
 CACNDLKFA
 40 IGILDVNCQKYDKFCFLQNKPVLLILLRDPIDSLKSFINVRHQKNGFNEIL
 KIDINNTDF
 DKINDRIVYVHESNGCFNPDNTNQKFPSLESIKALSDTNHWMLMYNIRRNKT
 IEFRRFNKI
 IYIDMMDIVGDKTLFTLEKLSKILNFSSPDKNKIFYQQLYSPLTVLLPCI
 45 IKVNNKVKI
 FVSNRFSVKNIQIMENCIDITDKFKEIFHENLIIFCSKDHFDSLINNQTLY
 NVVLEYINK
 FLISLKKRINVEKNKEVKVDDVLDYFKKNISVAKSYKDILDEELVYIKQHR
 PDIVASWKY
 50 YQEFERMCKELDENNQNPSSLFSNQ

55

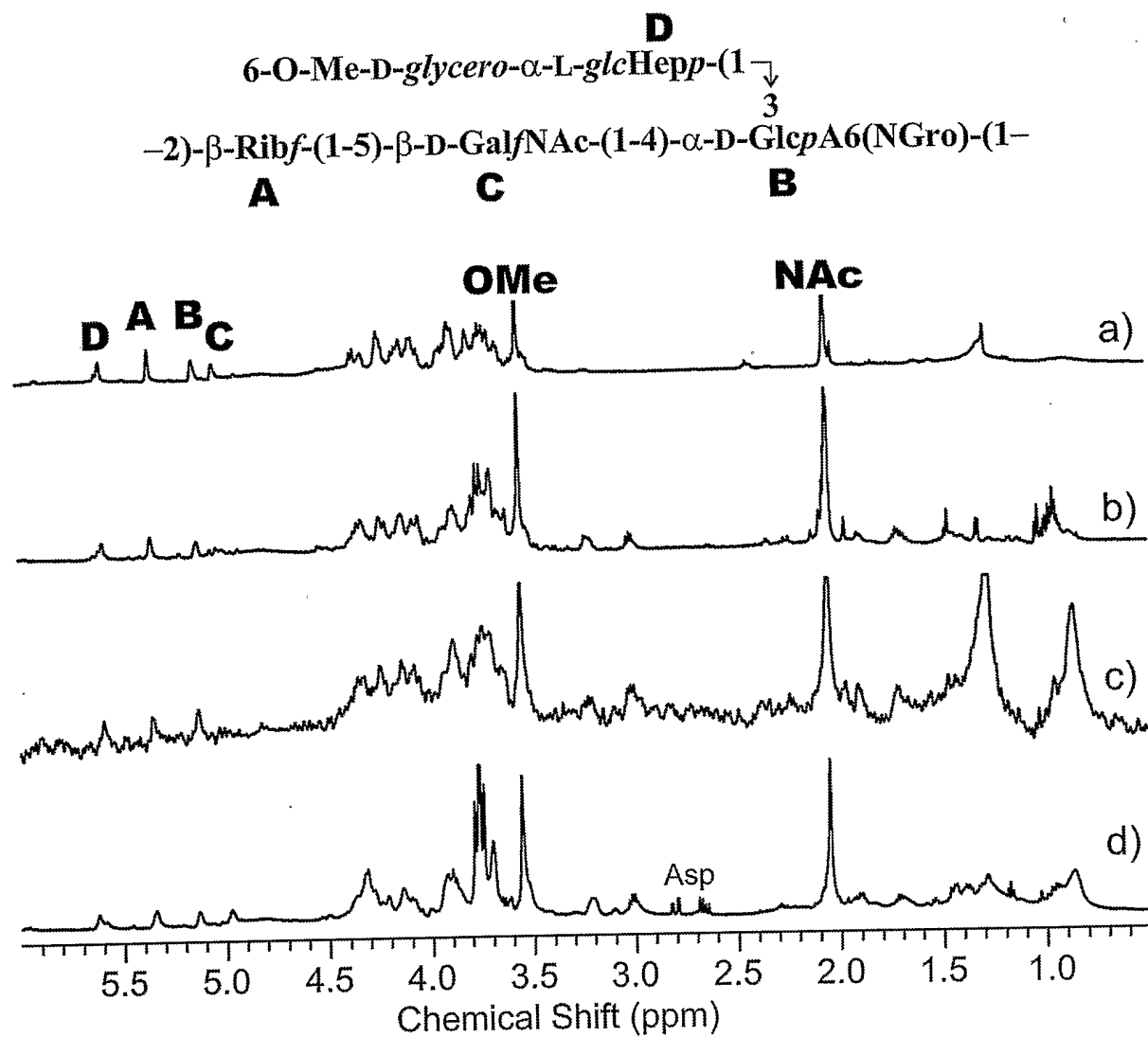
We Claim:

1. Use of the phosphoramidate $OP=O(NH_2)OMe$ or an immunologically active derivative thereof in the identification of *Campylobacter jejuni*.
- 5 2. The use of claim 1 wherein the phosphoramidate is used as a target for a binder.
3. The use of claim 2 wherein the binder is localized on the surface of a liposome.
4. The use of claim 2 wherein the binder is conjugated to a biologically
10 active molecule.
5. The use of claim 2 wherein the binder is conjugated to a marker radioisotope.
6. The use of claim 2 wherein the marker is a radioisotope.
7. The use of claim 2 wherein the marker is a fluorescent molecule.
- 15 8. The use of claim 2 wherein the binder comprises at least one bacteriophage or a portion thereof.
9. The use of any preceding claim wherein the binder comprises an antibody or fragment thereof.
10. The use of claim 1 wherein the phosphoramidate further includes an alkyl
20 group attached to the O which is attached to the P of the methyl amidophosphate group, wherein the alkyl group is a sugar producible in *Campylobacter jejuni*.
11. The use of claim 2 wherein the alkyl group is a naturally occurring sugar.
12. The use of claim 8 wherein the sugar is an enantiomer of a naturally
25 occurring sugar.
13. Use of alkyl methyl amidophosphate or an immunologically active derivative thereof in the identification of *Campylobacter jejuni*, wherein the alkyl group is a sugar producible in *Campylobacter jejuni*.
14. Use of the phosphoramidate $OP=O(NH_2)OMe$ or an immunologically active
30 derivative thereof as a vaccine in a mammal.
15. A method of modulating the adhesion of *C. jejuni* cells to a surface, said method comprising modulating the concentration of binders in surrounding fluid.

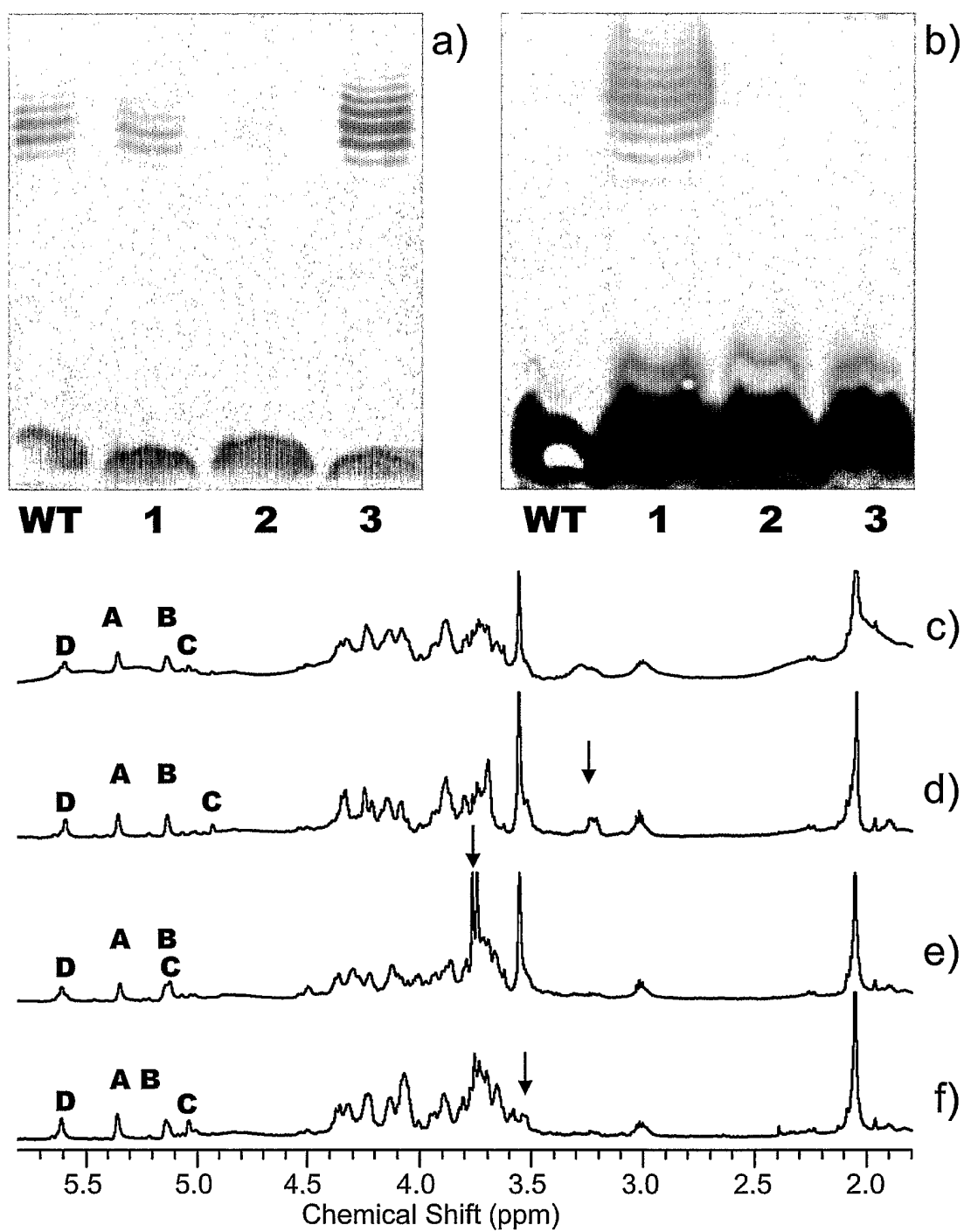
16. A phosphoramidate $OP=O(NH_2)OMe$ or an immunologically active derivative thereof.
17. A substantially pure alkyl methyl amidophosphate or an immunologically active derivative thereof.
- 5 18. A substantially pure pharmaceutical composition comprising the phosphoramidate of claim 16 or 17 and a physiologically acceptable carrier.
19. The composition of claim 18 including an immunogenic conjugate.
20. The composition of claim 18 comprising an immunostimulant.
21. The phosphoramidate of claim 16 or 17 linked through the O which is
10 attached to the P of the methyl amidophosphate group, to an amino acid or alkyl group wherein the alkyl group is a sugar producible in *Campylobacter jejuni* and the amino acid is an amino acid producible in *Campylobacter jejuni*.
22. *Campylobacter jejuni*.
- 15 23. A substantially pure binder which specifically binds the phosphoramidate of claim 16 or 17.
24. The binder of claim 23 which is an antibody or antibody fragment.
25. A kit comprising:
 - a) the binder of claim 23; and
 - 20 b) instructions for carrying out the method of claim 1 or 13.
26. Use of an isolated nucleic acid encoding *Campylobacter jejuni* Cj1421c or Cj1422c or a portion or variant thereof in producing a polypeptide having wild-type transferase activity the polypeptide producing an amino acid being useful in producing non-naturally occurring antigenic compounds
25 comprising the phosphoramidate of claim 16 or 17.
27. The use of claim 26 wherein the nucleic acid sequence encodes an amino acid sequence at least 90% identical to a wild type Cj1421c or Cj1422c sequence.
28. Use of an amino acid sequence encoding *Campylobacter jejuni* Cj1421c
30 or Cj1422c or a portion or variant thereof having wild-type transferase activity in producing an amino acid sequence useful in producing non-naturally occurring antigenic compounds comprising the phosphoramidate of claim 16 or 17.

29. The use of claim 27 wherein the amino acid sequence encodes a variant at least 90% identical to the wild type Cj1421c or Cj1422c sequence.

1/14

**FIG. 1**

2/14

**FIG. 2**

3/14

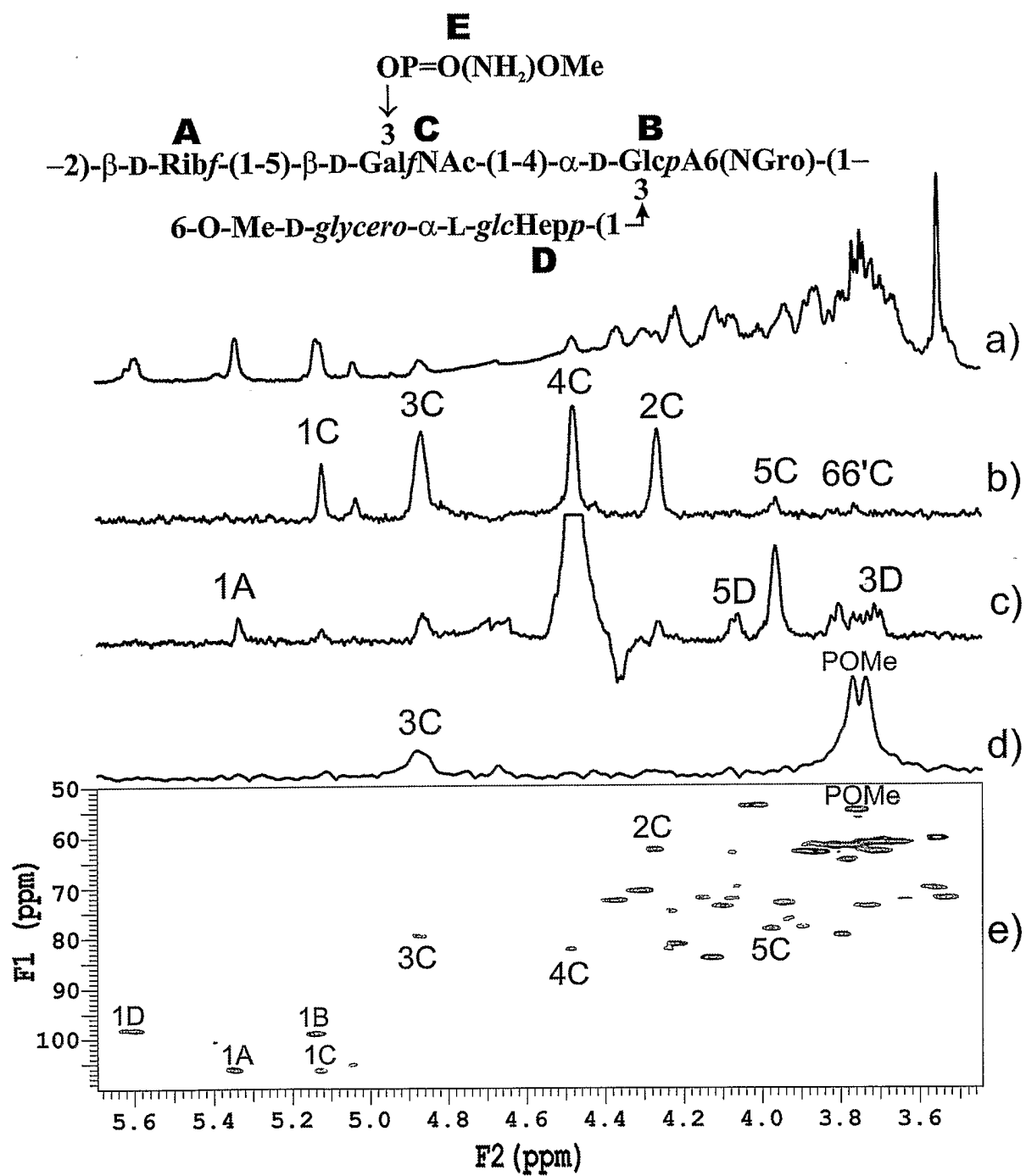


FIG. 3

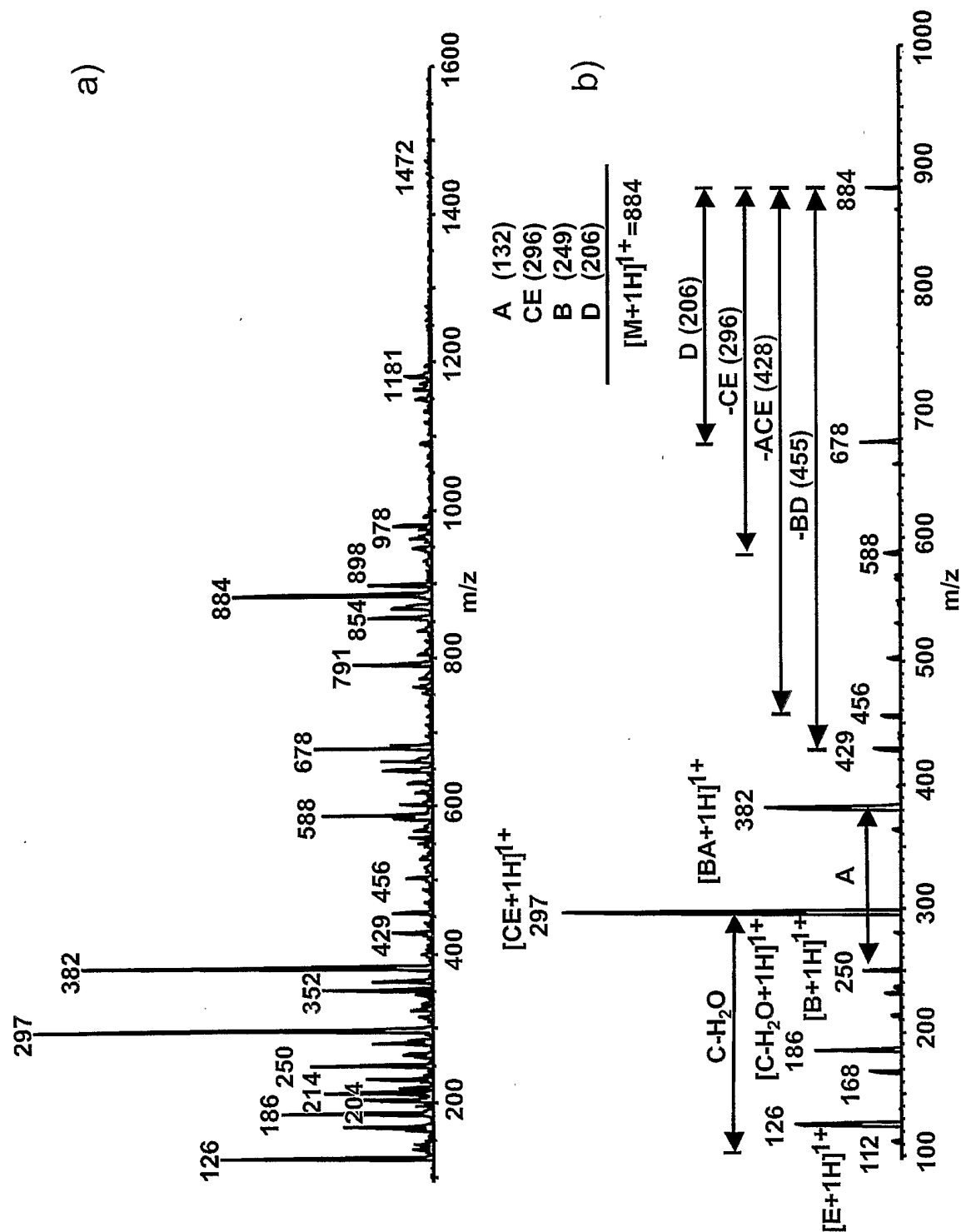


FIG. 4

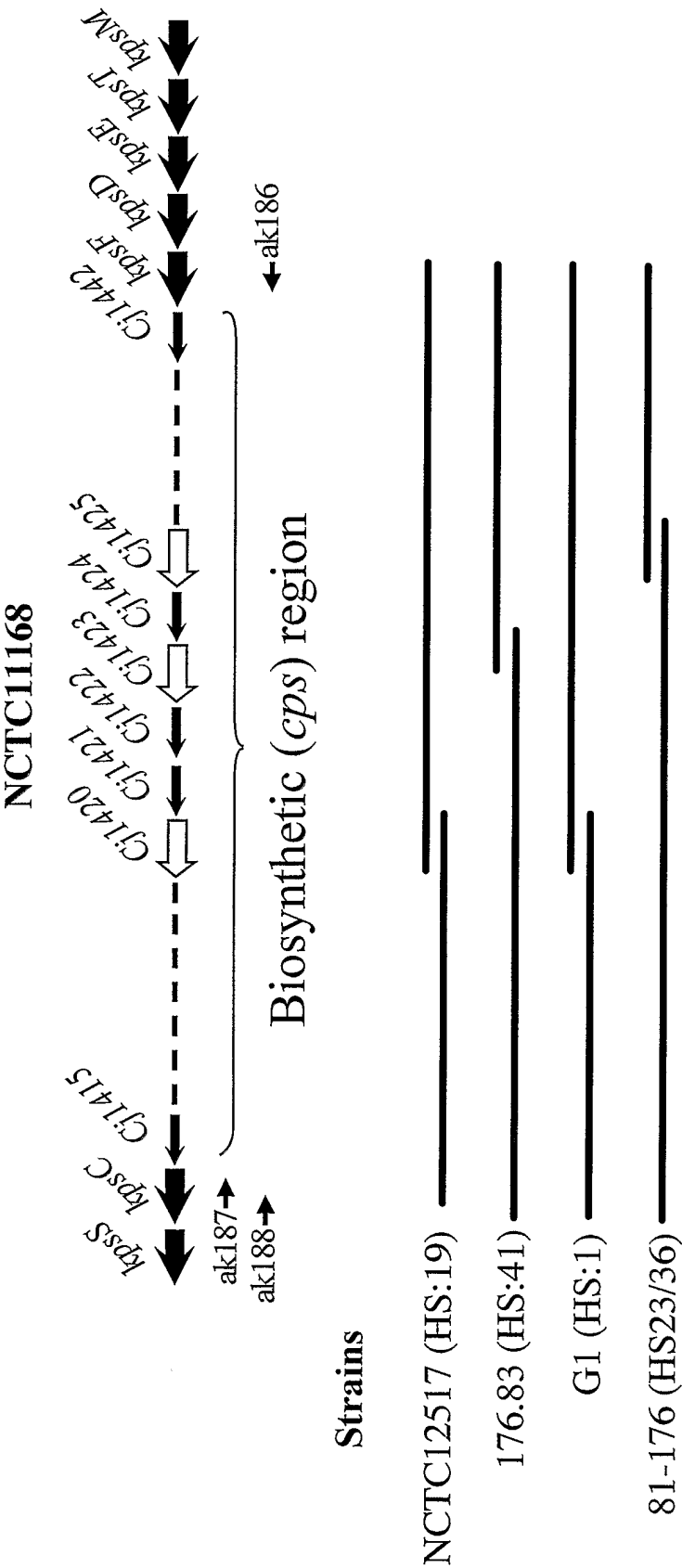


FIG. 5

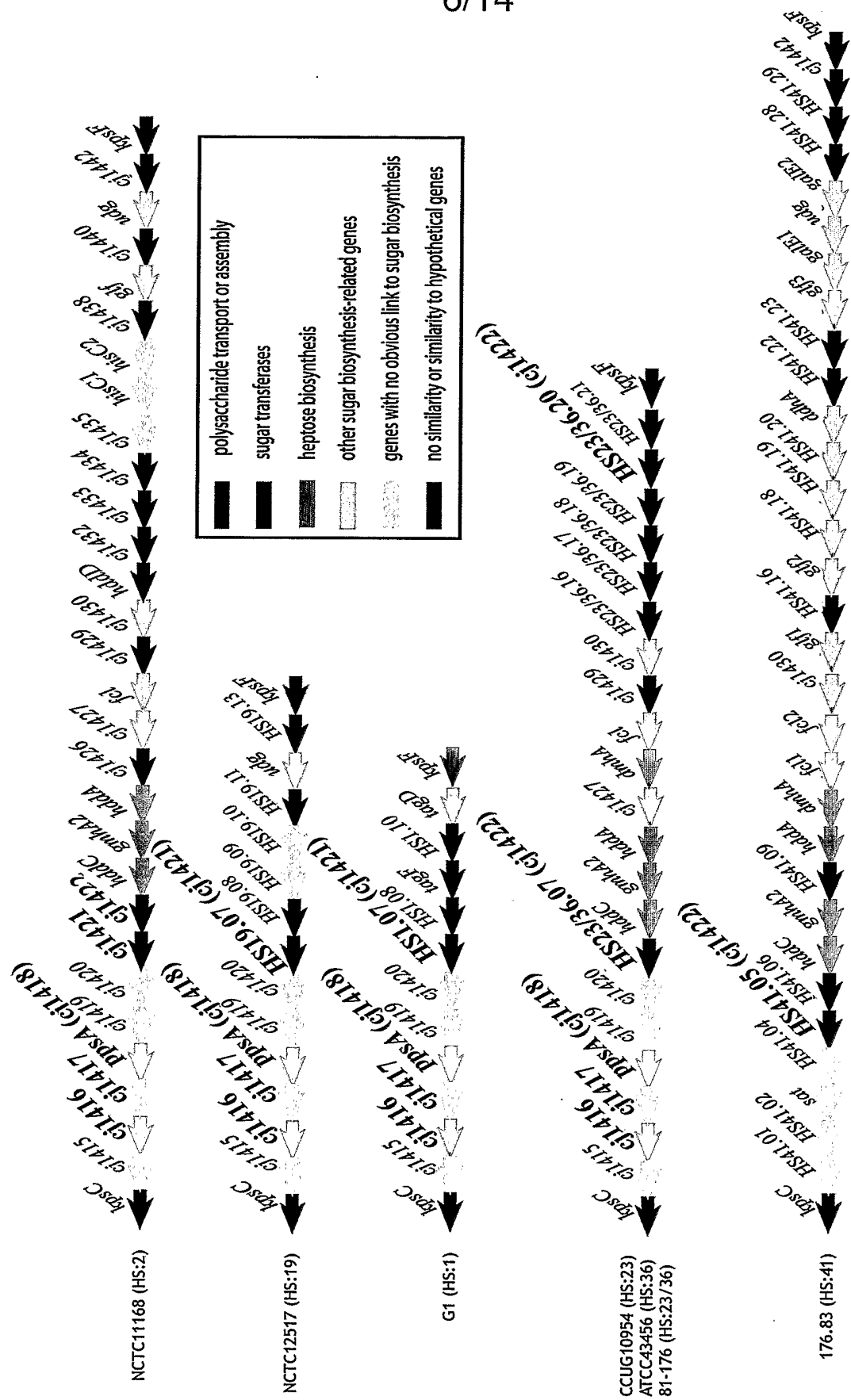


FIG. 6

7/14

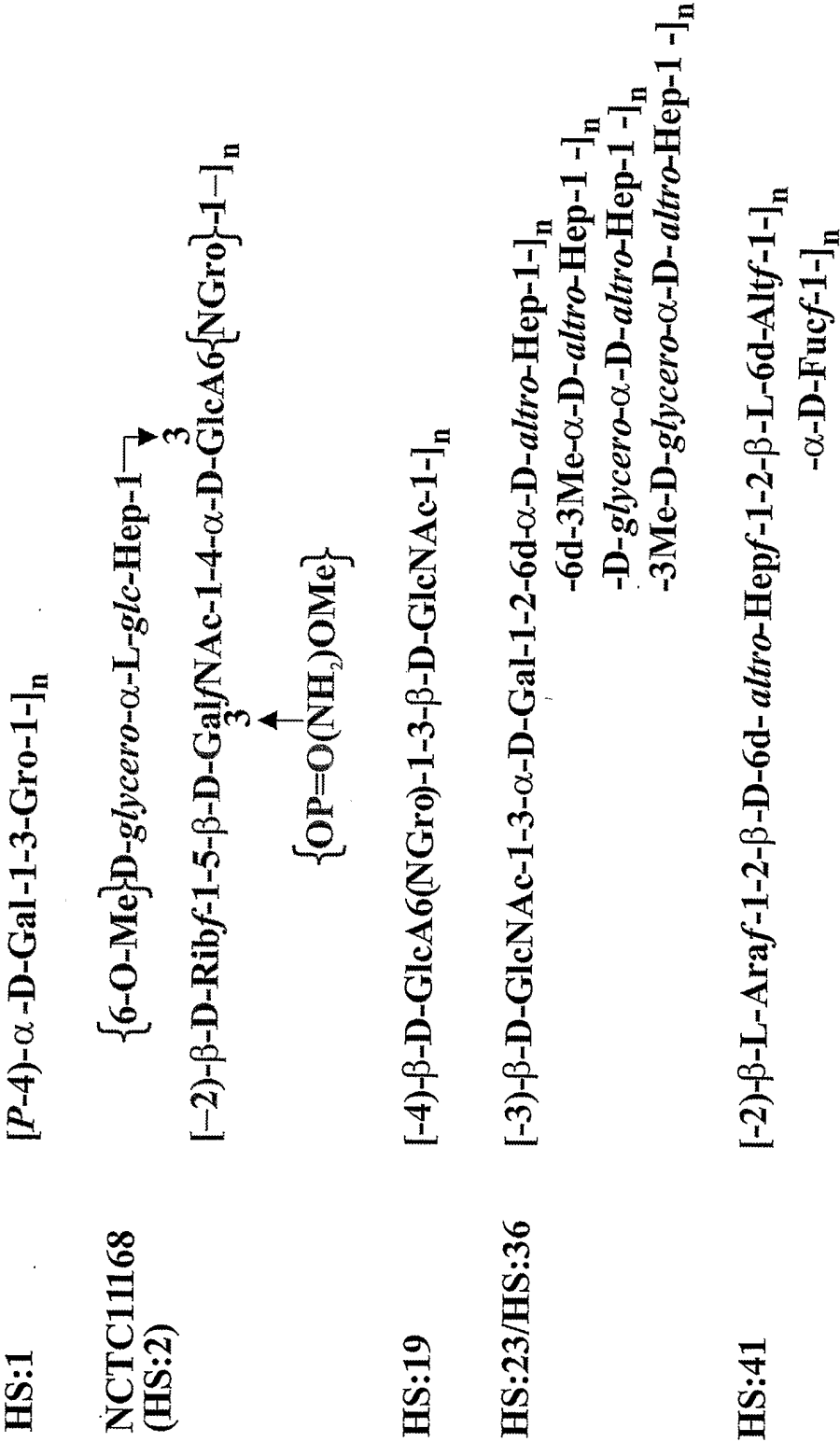
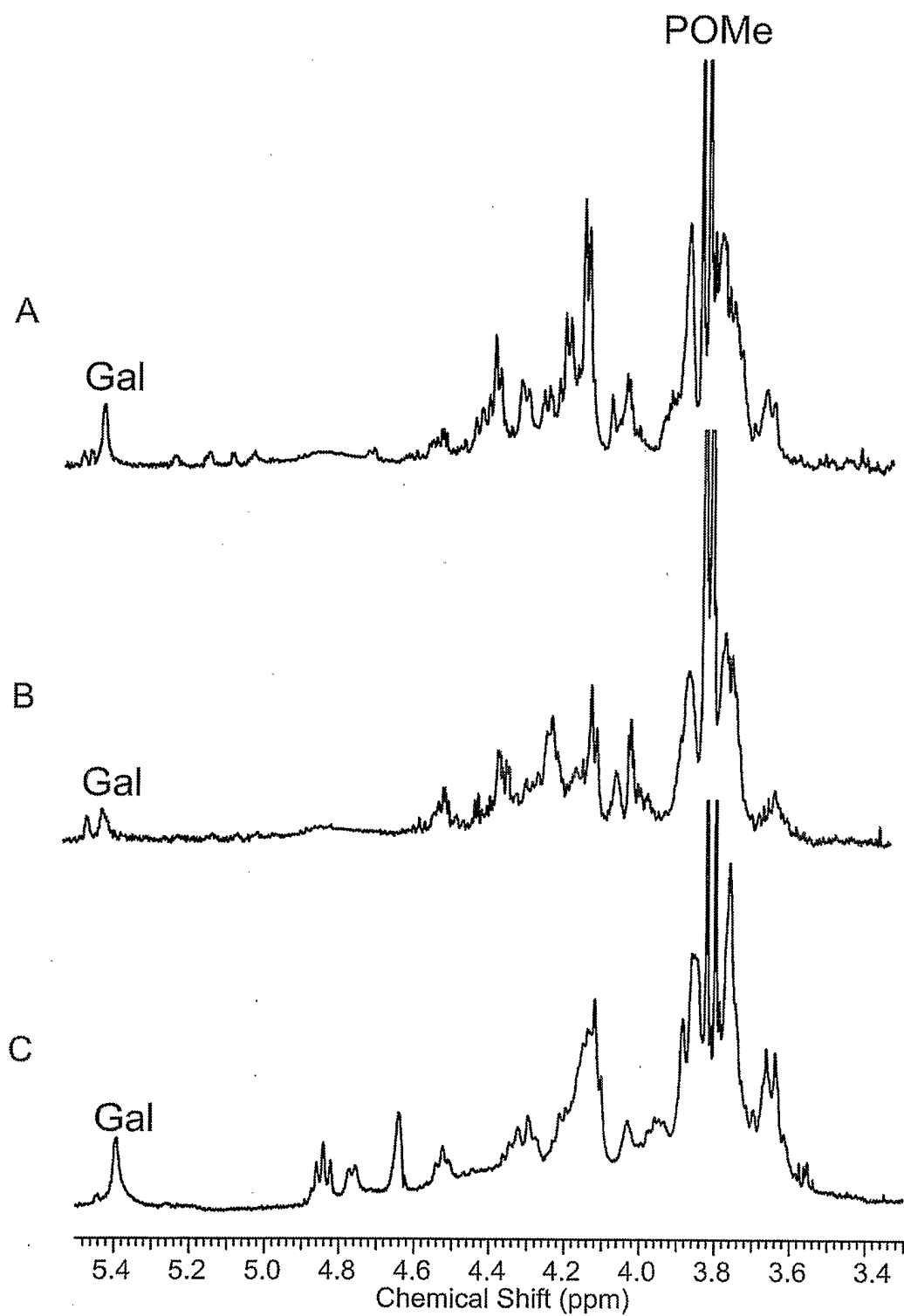
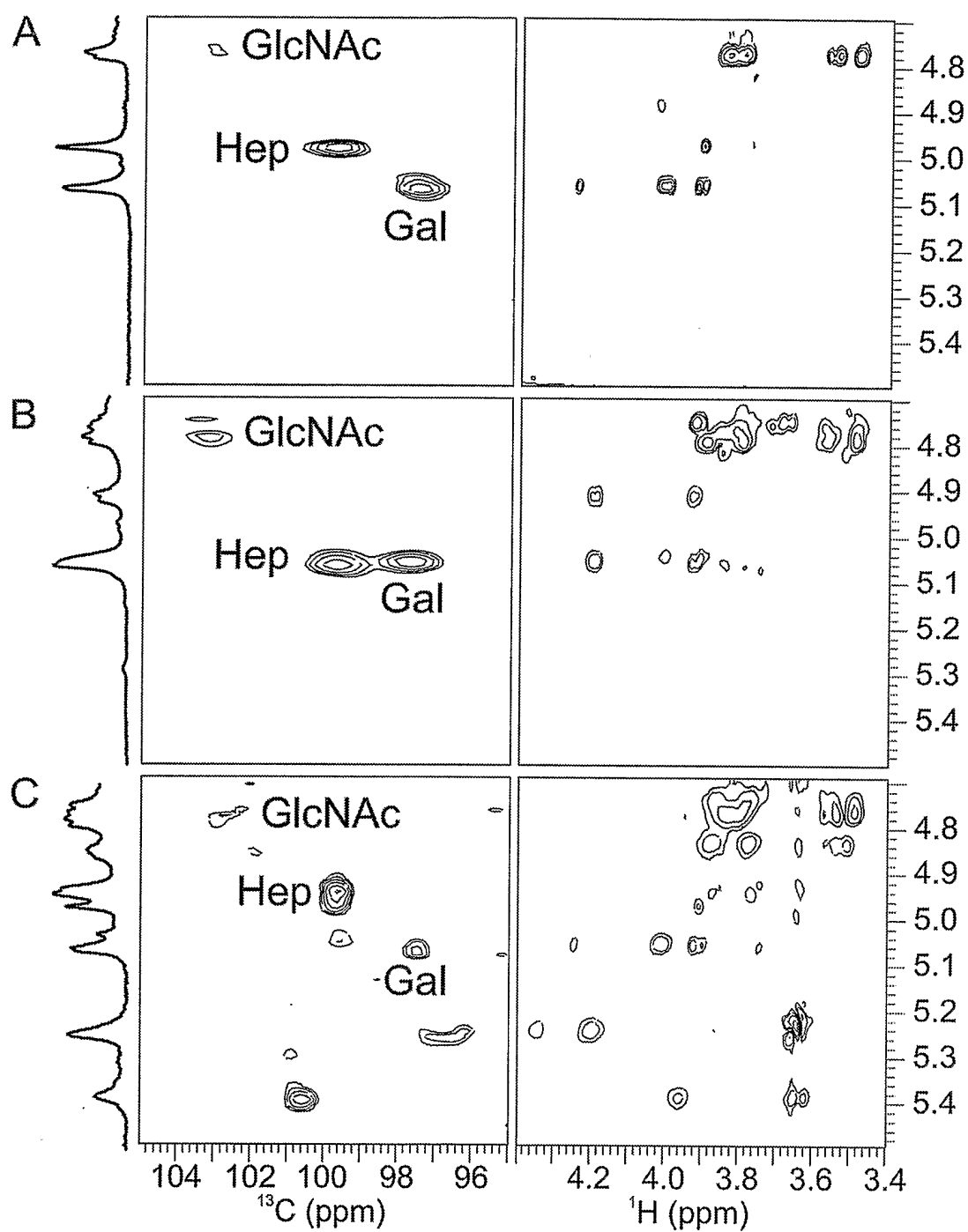


FIG. 7

8/14

**FIG. 8**

9/14

**FIG. 9**

10/14

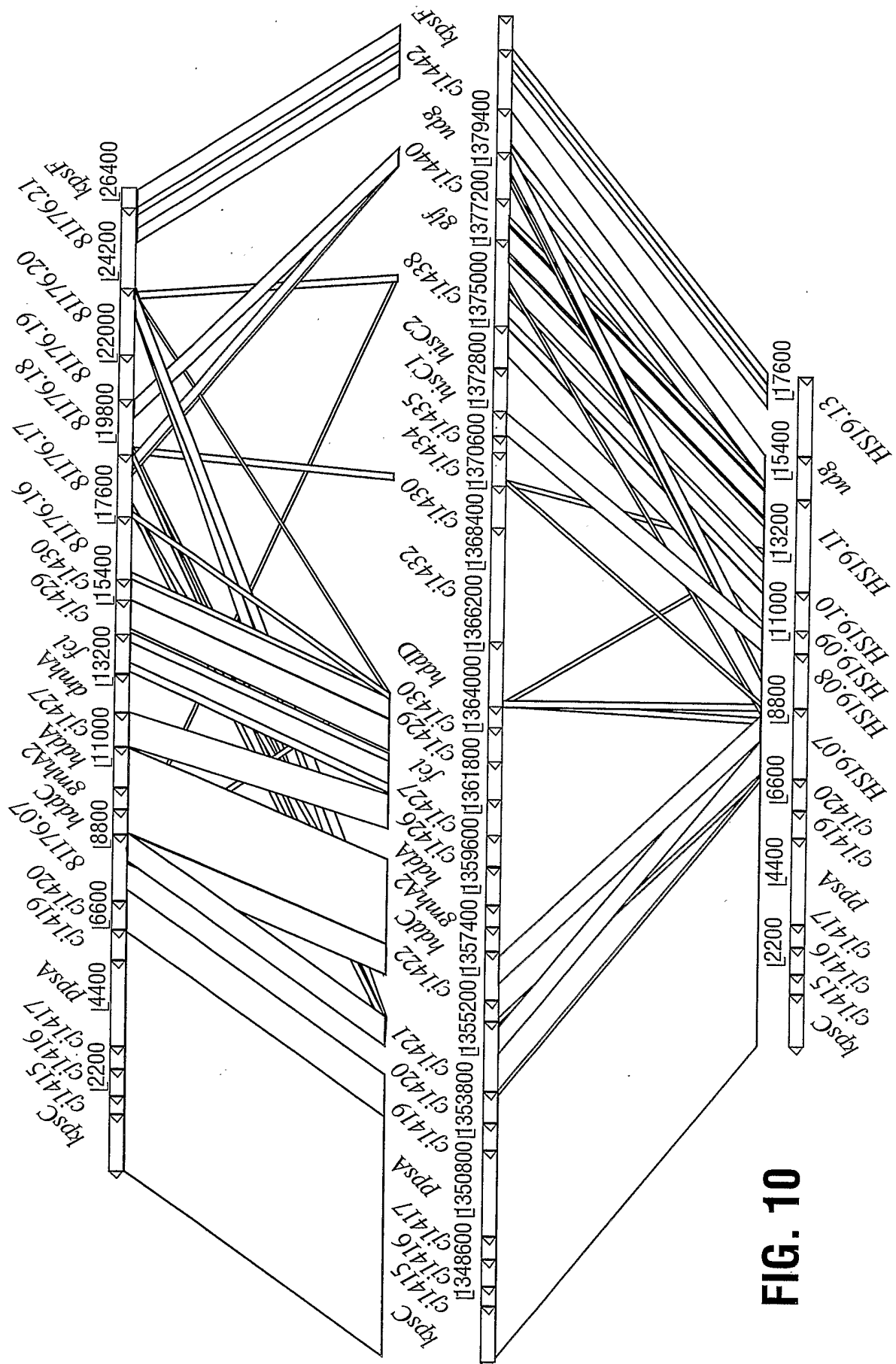


FIG. 10

11/14

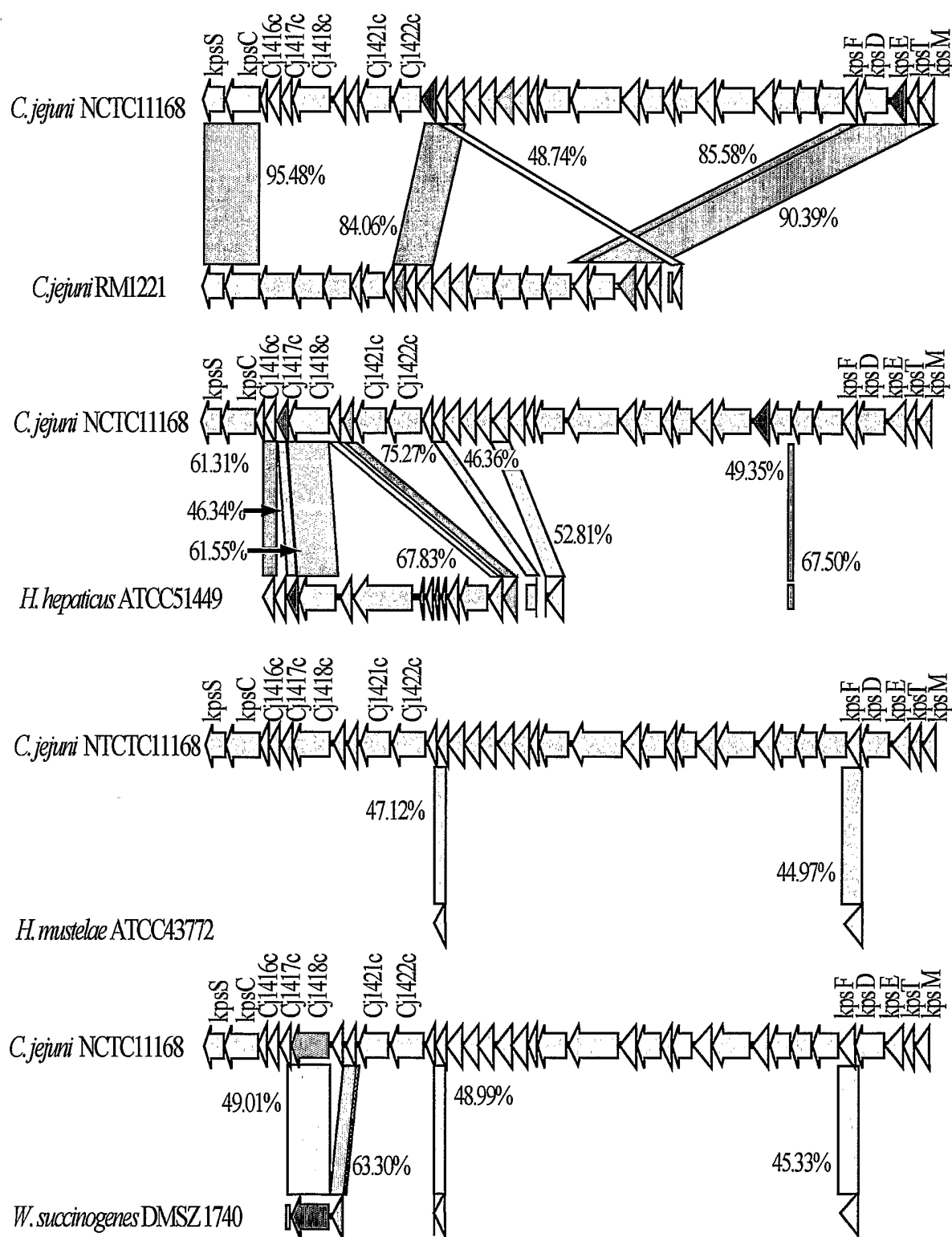
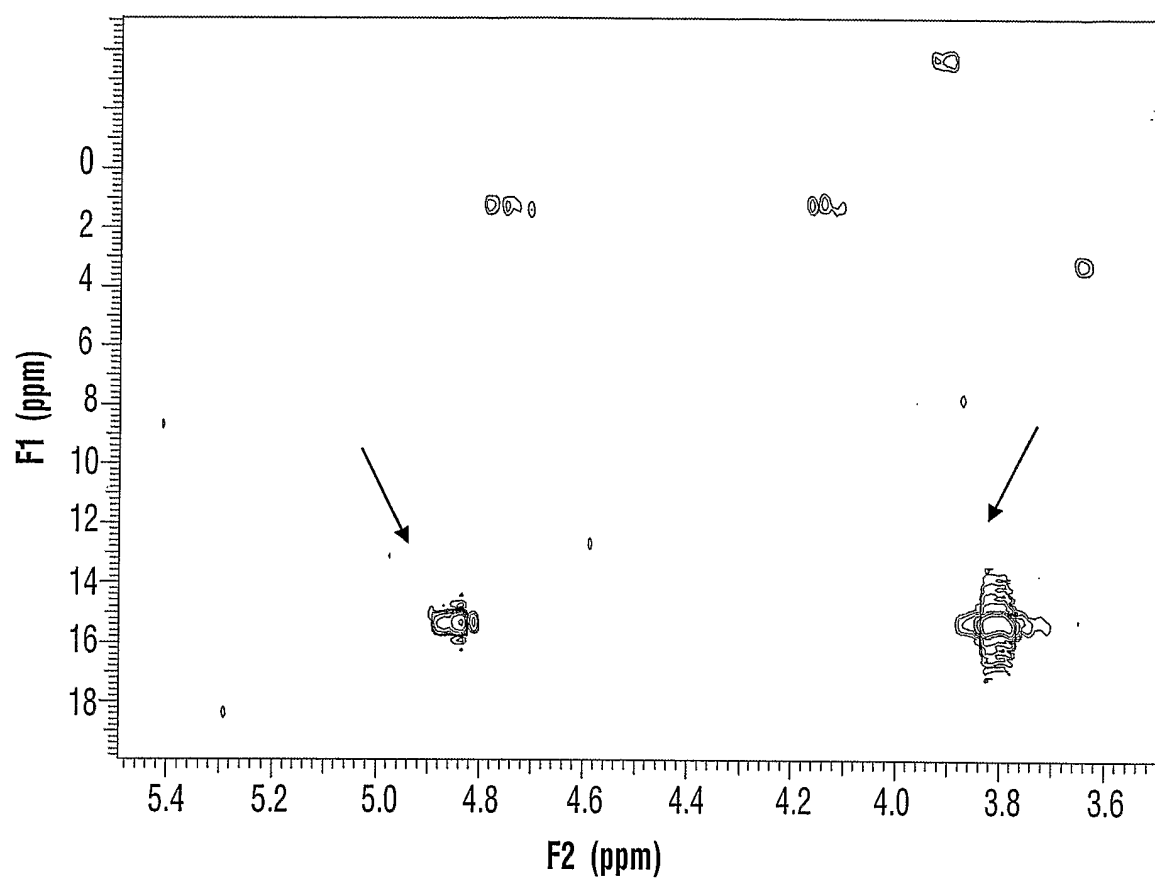
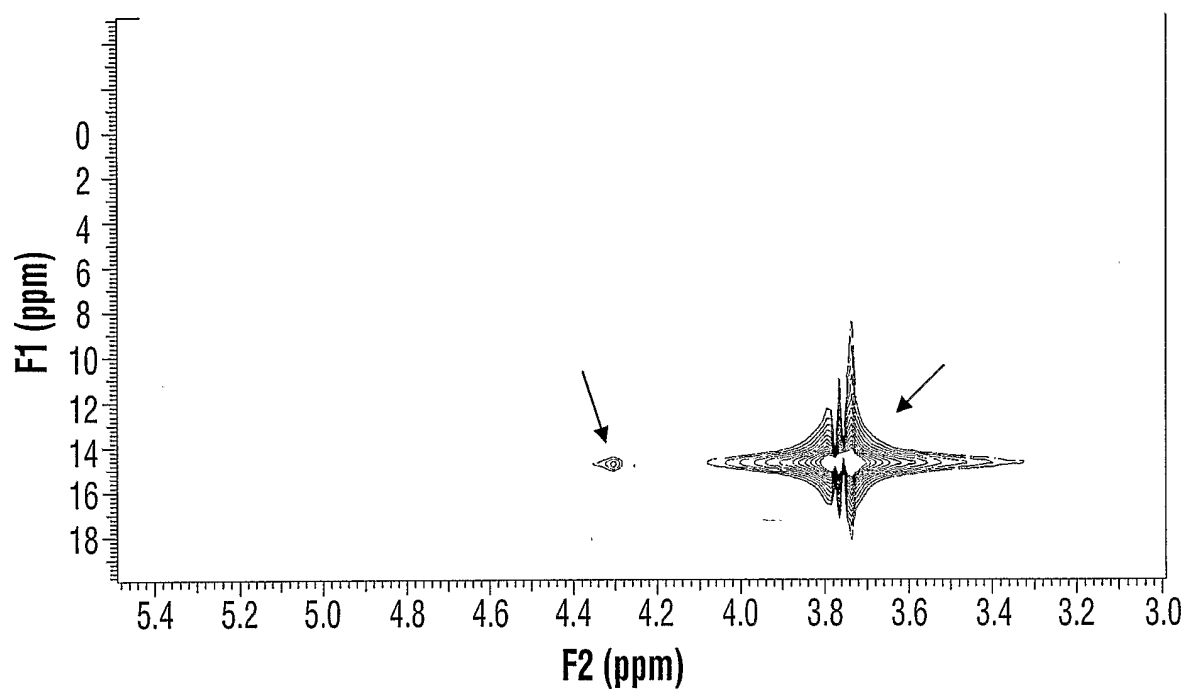
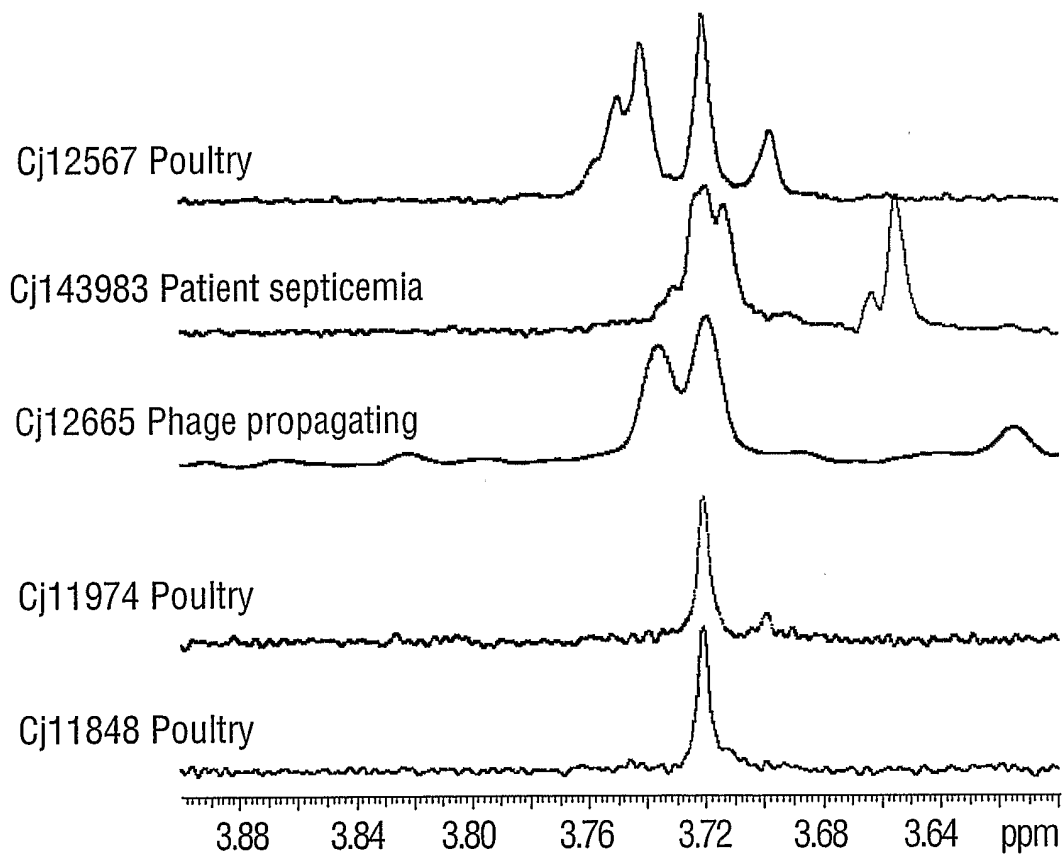


FIG. 11

12/14

**FIG. 12**

13/14

**FIG. 13****FIG. 14**

14/14

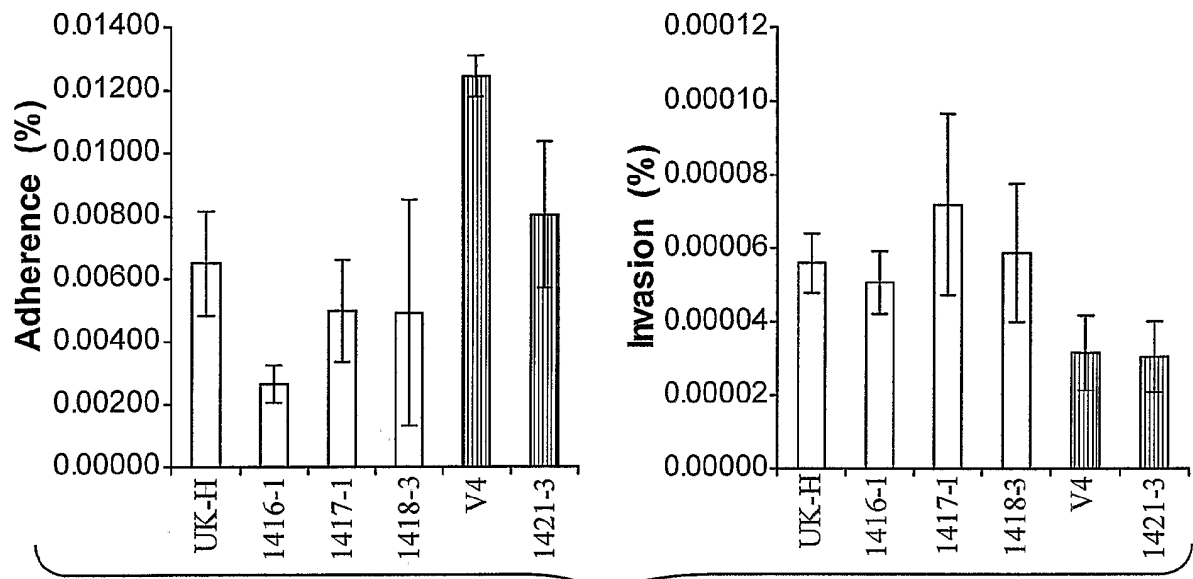


FIG. 15

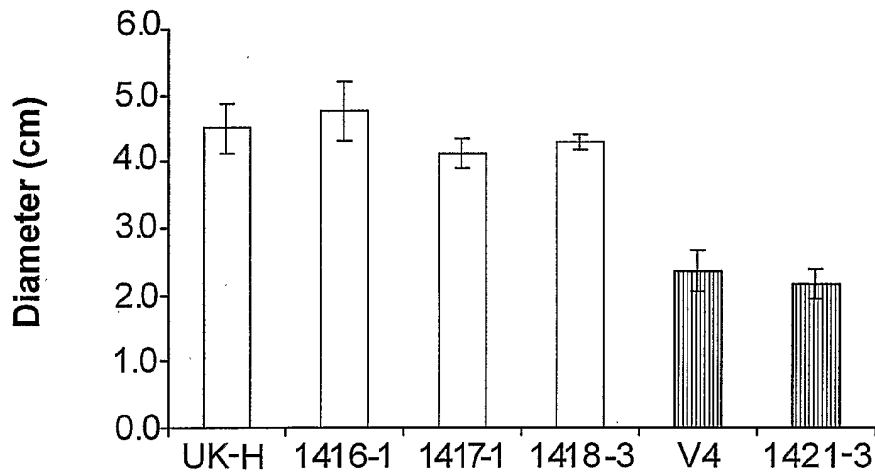


FIG. 16

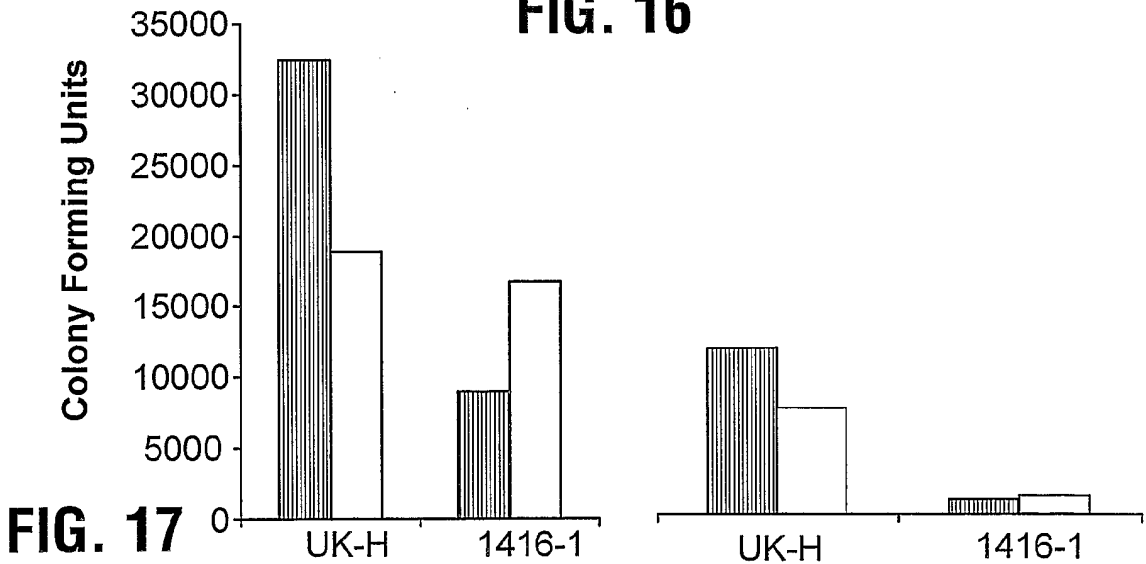


FIG. 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2004/001489

A. CLASSIFICATION OF SUBJECT MATTER

C12N1/20, C12Q1/04, A61K31/664, G01N33/569, C07K16/12, C07K14/205, C12N15/31, C12N9/10, C12N15/54, C12P21/02, A61K39/106, A61P31/04, C07F9/24, C07F9/22

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
C07F, C12Q 1/04, A61K 31/664, A61K 39/106; Chemical Abstracts

Electronic data base consulted during the international search (name of data base, and, where practicable, search terms used)
Canadian Patent Database, Delphion, Q-WEB, PubMed (NCBI), Internet (Google™ search engine)
Keywords: CH₃NO₃P, phosphoramidate, amidophosphate, Campylobacter

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	The Journal of Biological Chemistry vol. 278, no. 27, 4 July 2003 pages 24509-24520 SZYMANSKI, C. M. ET AL: "Detection of Conserved N-linked Glycans and Phase-variable Lipooligosaccharides and Capsules from Campylobacter Cells by Mass Spectrometry and High Resolution Magic Angle Spinning NMR Spectroscopy" see the whole document	16-18, 21
Y	---	19, 20
A	Eur. J. Biochem. vol. 269, no. 21, 2002 pages 5119-5136 ST. MICHAEL, F. ET AL: "The structures of the lipooligosaccharide and capsule polysaccharide of Campylobacter jejuni genome sequenced strain NCTC 11168" cited in the application see the whole document ---	1-14, 16-21
A	US, A, 5 200 344 (BLASER, M. B. ET AL) 6 April 1993 see the whole document	1-14

Further documents are listed in the continuation of Box C.

Patent family members are listed in annex.

* Special categories of cited documents :	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international-type search
13 December 2004 (13-12-2004)

Date of mailing of the international-type search report
17 December 2004 (17-12-2004)

Name and mailing address of the ISA/
Commissioner of Patents
Canadian Patent Office - PCT
Ottawa/Gatineau K1A 0C9
Facsimile No. 1-819-953-9358

Authorized officer
Katrina Campsall (819) 956-9874

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2004/001489

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons :

1. ☐ Claims Nos. :
because they relate to subject matter not required to be searched by this Authority; namely:

2. ☐ Claims Nos. :
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically :

3. ☐ Claims Nos. :
because they are dependant claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observation where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows :

1. Claims: 1-14, 16-21 (completely); 25 (partially)
Invention 1 concerns phosphoramidate and uses thereof, as well as, a composition comprising said phosphoramidate.
 2. Claims: 15, 23, 24 (completely); 25 (partially)
Invention 2 concerns a binder which binds to phosphoramidate, a kit comprising said binder and the use of binders in a method to modulate adhesion of *C. jejuni*.
 3. Claim: 22
Invention 3 concerns *Campylobacter jejuni*.
 4. Claims: 26-29
Invention 4 concerns the use of sequences encoding *Campylobacter jejuni* Cj1421c or Cj1422c.
-
1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
 2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
 3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos. :

 4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos. :

1-14, 16-21

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2004/001489

PATENT DOCUMENT CITED IN SEARCH REPORT	PUBLICATION DATE	PATENT FAMILY MEMBER(S)	PUBLICATION DATE
US5200344	06-04-1993	AU663858 B2	26-10-1995
		AU681043 B2	14-08-1997
		AU3042995 A	14-12-1995
		AU9071791 A	11-06-1992
		CA2096016 A1	14-05-1992
		EP0557467 A1	01-09-1993
		JP6505024 T	09-06-1994
		US5470958 A	28-11-1995
		WO9208485 A1	29-05-1992